The Black Hole Information Paradox

Thesis · September 2015
DOI: 10.13140/RG.2.2.22562.07363

Charles W. Robson
Tampere University

Imperial College London

Master's Thesis

The Black Hole Information Paradox

Author:
Charles Robson

Supervisor:
Professor Kellogg Stelle

Submitted in partial fulfilment of the requirements for the degree of Master of Science of Imperial College London

Theoretical Physics Group
Department of Physics, Imperial College London

September 2015

"How wonderful that we have met with a paradox. Now we have some hope of making progress."

Niels Bohr

"Nature is wont to hide herself."

Heraclitus

Abstract

A review of the black hole information paradox and its potential solutions is presented. Firstly we take a brief look into the history of black holes, some of the most useful mathematical tools used to investigate them, and the discovery that they are thermodynamical systems. A derivation of Hawking radiation using the Unruh effect is then presented before introducing information theory and a survey of modern attempts to resolve the information paradox. We conclude that the "fuzzball" proposal originating from string theory is the most promising solution to have been put forward so far. The recent proposal by Stephen Hawking involving supertranslations of the event horizon is also included.

Acknowledgements

I would like to thank Professor Stelle for agreeing to supervise this MSc project and for his suggestions regarding the recent work applying supertranslations to the paradox.

I also wish to thank my parents for their support.

Contents

List of Figures

Abbreviations

BHIP	**B**lack **H**ole **I**nformation **P**aradox
IEF	**I**ngoing **E**ddington-**F**inkelstein
OEF	**O**utgoing **E**ddington-**F**inkelstein
RN	**R**eissner-**N**ördstrom
AdS/CFT	**A**nti-**d**e **S**itter/**C**onformal **F**ield **T**heory

Chapter 1

Introduction

The black hole information paradox may be considered later in the century to be one of the defining collisions of the principles of modern physics.

Paradoxes have a long and fascinating history in the natural sciences and mathematics, they usually result from the use of concepts that are ill-defined or self-referencing, and quite often are not paradoxical at all (i.e. the "twin paradox" of special relativity) [1, 2]. They can also emerge from a clash of theories. The great physical paradox of the turn of the 20th century, the ultraviolet catastrophe, emanated from a clash of physical theories – between classical statistical thermodynamics and field theory [3]. This ultimately led to the genesis of quantum theory. In an analogous way, Albert Einstein's early gedankenexperiments exploited the discord between Galilean relativity and classical electrodynamics, leading to the special theory of relativity [4]. In each of these cases the conflict of ideas and their apparent discontinuities provided the conditions out of which new and deep ideas emerged.

The black hole information paradox is the result of a contradiction between the foundations of general relativity and quantum mechanics and has been called both a "serious crisis" [3], and "probably the most important issue for fundamental physics today" [5]. Before discussing the paradox in more detail let's take a look at a brief history of black holes.

A black hole is an object so massive that the gravitational field it produces prevents even light escaping its pull. The size of the black hole is defined by the size of its *event* *horizon*, anything that falls past the event horizon can never escape. Except possibly information, as we will see later.

The first hint that black holes may be lurking in the universe occurred during the late eighteenth century when the Reverend John Michell, using Newton's corpuscular theory of light, found that if a star were sufficiently massive, its escape velocity would exceed the speed of light (which was known with remarkable accuracy at the time) [6, 7]. Around the same time as Michell's work, the eminent French scholar Pierre-Simon Laplace discovered the same result and put it on a more mathematical grounding [8]. By the 19th century the corpuscular theory favoured by Newton [9] had gone out of fashion, replaced by the wave theory of light (based on work by Christiaan Huygens, Augustin-Jean Fresnel and Thomas Young amongst others) [10, 11], after which time the idea of dark stars was left as a theoretical pecularity.

It took until the publication of Einstein's general theory of relativity and Karl Schwarzschild's elegant vacuum metric solution that a black hole geometry could be derived and its features – such as its event horizon – parameterized and studied [12]. Despite this, Einstein and others refused to believe in the existence of black holes, believing them to be just too exotic, and produced work attempting to show how massive objects (such as stars) could never implode to such a degree that they would produce black holes [13].

By the 1920s high-density white dwarf stars had been discovered and an upper limit to their mass – the Chandrasekhar limit – had been theorised, past which the star would continue to collapse with gravitational effects overcoming the electron degeneracy pressure [14]. It was natural to consider what would happen to a star with a mass that exceeded this limit and even denser neutron stars were studied in the 1930s leading to a further implosion limit, the Tolman-Oppenheimer-Volkoff limit [15]. In 1939 work by Oppenheimer, Volkoff and Snyder [16, 17], utilising Einstein's field equations and Schwarzschild's vacuum solution, suggested that if a ball of gas were sufficiently massive (and suffered from no outward pressure to counteract collapse) the gas would implode indefinitely, producing an inifinitely dense singularity and an event horizon. This paved the way for further investigations and for faith in the existence of black holes by the scientific community to grow substantially by the 1960s, largely thanks to improved astrophysical computer simulations of collapsing matter; from this time up until the middle of the 1970s a "golden age" of classical black hole physics was underway [10].

Work done at this time led to more features of black holes being predicted and studied, including their mass, electrical charge and angular momentum, leading to the "no-hair theorem", this stated that the information concerning the nature of a collapsing body that formed a hole had to be encoded purely in terms of these three features – an external observer cannot see any other features of the collapsing mass due to the presence of the event horizon [10]. Another important discovery treated the possibility of extracting work from a rotating black hole, theorised by Penrose and Floyd [18], as well as the cosmic censorship hypothesis which posited that singularities should not be visible in the universe – they should always be hidden behind event horizons [19].

At the start of the 1970s quantum theory began to be combined with the classical general relativistic treatment of black holes, during this time the four classical laws of thermodynamics were found to be strikingly analogous to the rules that had been developed to describe black holes [20]. Most intriguingly the generalised second law suggested by Bekenstein [21] gave an entropy for a black hole proportional to its horizon area. Stephen Hawking's discovery in 1974 [22] that black holes radiate was a revolutionary discovery which cemented the relationship between black hole physics and thermodynamics and uncovered a deep, mysterious link between quantum gravity and thermodynamics that is still being studied extensively today [10, 23].

Stephen Hawking's papers [24, 25] on the emission of radiation from black holes suggested that the information inherent in a system that fell into a black hole would be forever lost, radiated away as black body radiation until eventually the black hole evaporated away completely, a conclusion incompatible with quantum theory. Around this time several other papers were published which highlighted a possible deep link between black holes and thermodynamical systems; firstly an analogue second law was put forward which associated the entropy of a black hole with its area, and then three other laws of black holes were found with striking similarities to the zeroth, first, and third laws of thermodynamics [20–22].

The destruction of information, as implied by the thermal nature of Hawking radiation, and also the origin of the extremely large entropy of a black hole (seemingly contradicting the "no-hair theorem") are puzzles that demand – and stimulate the search for – a full theory of quantum gravity. At this time there is still no consensus regarding what a theory of quantum gravity should be, however string theory is one of the most promising and most studied candidates, and its combination with supersymmetry has led to a strikingly accurate rederivation of the entropy of certain types of black holes (which we shall look into in more detail later on) [10, 26].

Although some prominent physicists [27] were content to give up the tenet of information conservation, other theorists [28] thought this conclusion to be deeply unsatisfactory and a step too far as it contradicted a foundational principle of quantum mechanics, namely that the evolution of a quantum state should be unambiguously determined by an invertible unitary operator [29]. Many ideas were put forward after the discovery of Hawking radiation that attempted to come up with ways of somehow conserving information during black hole evaporation, these included the information being somehow encoded within the Hawking radiation, being contained in Planck-scale "remnants", leaking out right at the end of evaporation, being encoded within distortions of the event horizon and even the possibility of the information being stored in baby universes apart from our own [30].

The AdS/CFT correspondence [31], which states that there exists a duality between gravitational theories in the bulk and quantum conformal field theories on the boundary, convinced many physicists that information had to be conserved in black hole evaporation somehow and that all that was needed to be done was to find the mechanism [32–35].

In the early 1990s a promising solution appeared in the form of "complementarity" [3, 36], which posited that information falling into a black hole both passes through the event horizon *and* bounces back off of it to be collected by an external observer. This counterintuitive idea is not contradictory as *no single observer* can detect both the inward and outwardly propagating information. However the conceptual difficulties remained and in 2013 the AMPS "firewall" paradox [37] showed that complementarity led to a very strange conclusion, that there must be a layer of high-energy quanta – a firewall – at the event horizon that would fry any observer who fell in. The equivalence principle of general relativity suggests (ignoring tidal effects) that an observer should feel nothing strange when falling through the horizon and so the firewall conclusion has proven to be very perplexing indeed.

An exciting solution from string theory known as the "fuzzball" proposal describes the microstates of black holes as bound states of strings and branes, effectively removing the traditional event horizon and singularity; fuzzballs allow information to leak out in Hawking radiation and can be used to correctly rederive the Bekenstein entropy of the hole, a very suggestive result [26].

We shall begin by briefly looking at the basic properties of black holes and some of the mathematics that is commonly utilised when studying them. Then the treatment of black holes as thermodynamical systems will be introduced before moving on to a derivation of Hawking radiation via the derivation of Unruh radiation – experienced by an eternally accelerating observer in empty space. This will be followed by a short discussion of information theory and a precise statement of the information paradox. The final part of the work will be a review and appraisal of the main solutions to the paradox that have been put forward so far.

Chapter 2

Black Holes

2.1 The Basic Properties of Black Holes

In this section I have mainly followed the treatments given in *Dowker* [12], *Carroll* [15], *Wald* [38], and *Townsend* [39]. This section presents a small sample of the most basic features of black holes, the most important part in relation to the information paradox is the discussion of the Rindler metric.

The study of black holes in general relativity is done using a wide variety of coordinate systems (and sometimes without any at all [40]). Some coordinate systems have *singularities*, places where the spacetime point is mathematically ill-defined. Singularities can occur simply due to an artifact of the system being used, however they can also result from a physical property of the spacetime being considered, i.e. as a result of the curvature associated with a spacetime point being infinite.

One of the earliest and most important coordinate systems that was used to study black holes was the Schwarzschild system mentioned earlier. This solution is valid in a vacuum region surrounding a spherically symmetric distribution of mass and can be derived from Einstein's field equations [15]. The Schwarzschild line element is [12]:

$$ds^2 = -\left(1 - \frac{2GM}{rc^2}\right) c^2 dt^2 + \left(1 - \frac{2GM}{rc^2}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \qquad (2.1)$$

where r is a radial coordinate which runs from zero to infinity, θ the polar angle, ϕ the azimuthal angle, M the mass of the black hole, and t the time coordinate. We will see later how these coordinates should not necessarily be taken as literally representative of physical quantities. The coordinate systems we will look at contain an infinity corresponding to the physical singularity at the centre of black holes. These singularities can be seen to be physical and not an artifact of the mathematical system in use by finding an invariant measure of the curvature of spacetime such as the Kretschmann scalar, this measure is formed from two Riemann curvature tensors:

$$K = R_{abcd}R^{abcd}, \tag{2.2}$$

for a Schwarzschild black hole its value is $K = \frac{48G^2M^2}{c^4r^6}$ [41]. This can be seen to diverge as r reaches zero which is expected as the centre of the hole is normally thought of as having infinite density.

From a quick inspection of (2.1) it is apparent that there are two singularities. One at $r = 0$ and the other at $r = \frac{2GM}{c^2}$, the latter known as the Schwarzschild radius (or event horizon radius). In the subsequent analysis we will see that the Schwarzschild radius is not a physical singularity but only a feature of the metric. Due to Birkhoff's theorem [15], the Schwarzschild metric is the *unique* spherically symmetric solution to the vacuum field equations of Einstein.

Einstein's field equations are (with zero cosmological constant) [15]:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}. \tag{2.3}$$

By multiplying both sides of the field equations by $g^{\mu\nu}$, and in addition using $g^{\mu\nu}R_{\mu\nu} = R$, $g^{\mu\nu}T_{\mu\nu} = T$ and $g^{\mu\nu}g_{\mu\nu} = D$, where D is the number of spacetime dimensions, we obtain:

$$R - \frac{D}{2}R = \frac{8\pi G}{c^4}T. \tag{2.4}$$

Solving for R, inserting the solution into (2.3) and setting $D = 4$, the field equations become:

$$R_{\mu\nu} = \frac{8\pi G}{c^4}\left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}\right). \tag{2.5}$$

The form of the field equations given in equation (2.5) clarifies the effect of setting the stress-energy tensor $T_{\mu\nu}$ to zero, namely it forces the Ricci tensor $R_{\mu\nu}$ to also vanish.

For a simple way to study the dynamics of black holes we shall look at an imploding spherically symmetric ball of pressureless dust. Due to Birkhoff's theorem we know that the Schwarzschild metric must describe the region outside of the ball (and due to continuity of the metric must also apply to the surface). For a particle constrained to the surface of the ball and moving purely in a radial direction we can analyse its motion as it falls towards $r = 0$.

The action of this contrained particle is [12]:

$$S = \frac{1}{2m} \int \left(g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} - m^2 \right) d\tau, \tag{2.6}$$

where τ is the proper time along the particle's worldline and m is its mass.

The four-position of the particle can be parameterized as $x^\mu(\tau) = (t(\tau), r(\tau), \theta_0, \phi_0)$; θ_0 and ϕ_0 are both constants due to the motion being purely radial; from now on for ease of notation we shall assume $c = G = 1$ unless otherwise stated [12].

Defining the radius of the collapsing sphere (and therefore the position of the particle) as $R(t)$, equivalent to $r(\tau(t))$, and solving the Euler-Lagrange equations for the action produces a conserved quantity:

$$\epsilon = \left(1 - \frac{2M}{R} \right) \frac{dt}{d\tau} \tag{2.7}$$

which can be interpreted as the particle's energy per unit rest mass, measured by an inertial observer positioned at an infinite radial distance [12].

It can be derived from the Euler-Lagrange equations that $\frac{dR}{dt}$ vanishes at two values of R: one at $R = 2M$ and one at $R = \frac{2M}{1-\epsilon^2}$ (due to the fact that the particle is assumed to start its trajectory from rest at this latter point) [12]. But what is the correct interpretation of the apparent suspension of the particle at $R = 2M$ and why does the particle take an infinite time t to reach $R = 2M$ from its starting position? The reason for such a strange result stems from our use of t, which is only a valid measure of time for an observer stationed at infinity, replacing this parameter with the proper time τ of the infalling particle we find that the time taken for it to reach $R = 2M$ (and also $R = 0$) from its starting point is finite [12].

As we have seen we must be careful when interpreting coordinates in general relativity. From equation (2.1) it can be noted that the Schwarzschild metric coefficients change sign inside the event horizon, i.e. the r coordinate suddenly behaves like a time coordinate and t behaves like a radial coordinate; this makes our calculations seem slightly shaky. Next we will look at some other coordinate systems which are better behaved at the event horizon and allow for easier physical interpretations.

Eddington-Finkelstein coordinates are useful for studying the region across the event horizon of a black hole and are derived by firstly using *tortoise coordinates* (named memorably by John Wheeler) given by [42]:

$$r_* = r + 2M\ln\left(\frac{r-2M}{2M}\right). \tag{2.8}$$

Ingoing null geodesics followed by photons can be defined by $\nu =$ constant where $\nu = t + r_*$ and outgoing null geodesics by $u =$ constant where $u = t - r_*$ [12].

Ingoing Eddington-Finkelstein coordinates are defined by transforming the t coordinate to ν in the Schwarzschild metric to get:

$$ds^2 = -\left(1 - \frac{2M}{r}\right)d\nu^2 + 2d\nu dr + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2. \tag{2.9}$$

As can be easily seen there is no singularity in the IEF metric when $R = 2M$, and also the determinant of the metric is non-zero at this point, indicating that the metric is invertible and regular at the horizon [12].

What would the metric look like for an eternally accelerating observer? This observer exists in *Rindler space* (which is a subregion of Minkowski space) and we will now look at it in some detail as it is particularly useful in the study of black hole physics, especially in the derivation of Hawking radiation as we will see later. I have followed the treatments of Rindler space given in *Dowker* [12], *Carroll* [15], and *Blau* [43].

For simplicity let's analyse the physics of an accelerating observer in 1+1 dimensions. In the following derivations we take the Minkowski signature as $(-+)$. The observer is travellling in the x-direction and has an acceleration of magnitude α. The trajectory can be parameterized in Minkowski spacetime by [15]:

$$x^\mu(\tau) = \left(\frac{1}{\alpha}\sinh(\alpha\tau), \frac{1}{\alpha}\cosh(\alpha\tau)\right), \tag{2.10}$$

where τ is the observer's proper time. In Minkowski spacetime the acceleration two-vector is $a^\mu = \frac{d^2x^\mu}{d\tau^2}$; finding the components of the acceleration using equation (2.10) it is easy to see that $\sqrt{a^\mu a_\mu} = \alpha$, therefore the path parameterized above does indeed describe an observer eternally accelerating with magnitude α. Using the identity $\cosh^2\theta - \sinh^2\theta = 1$ [44] it is apparent that the Rindler path (path of constant acceleration) adheres to

$$x^2(\tau) = t^2(\tau) + \frac{1}{\alpha^2}, \tag{2.11}$$

x asymptotes to $-t$ at past null infinity and asymptotes to t at future null infinity.

Let's choose a new set of coordinates (η, ξ) that are better "*adapted to uniformly accelerated motion*" [15], obeying:

$$\begin{aligned} t &= \frac{1}{a}e^{a\xi}\sinh(a\eta), \\ x &= \frac{1}{a}e^{a\xi}\cosh(a\eta). \end{aligned} \tag{2.12}$$

Both η and ξ range from negative to positive infinity. Region I of Figure 2.1 is covered by $x > |t|$.

The Rindler path previously given by equation (2.10) is now given in terms of the new coordinates:

$$\begin{aligned} \eta(\tau) &= \frac{\alpha}{a}\tau, \\ \xi(\tau) &= \frac{1}{a}\ln\left(\frac{a}{\alpha}\right). \end{aligned} \tag{2.13}$$

Under the new coordinates, the Minkowski metric becomes the Rindler metric describing the frame of the eternally accelerating Rindler observer:

$$ds^2 = e^{2a\xi}\left(-d\eta^2 + d\xi^2\right). \tag{2.14}$$

The Rindler spacetime is embedded in Minkowski spacetime. Region I is the only region accessible to an observer with constant acceleration in the positive x-direction [15]. The lines of $x = \pm t$ are asymptotically approached by Rindler observers at past and future null infinity and so the lines demarcate types of horizon, we will see exactly what types of horizon in section 3.2 on the Unruh effect (strictly speaking these horizons

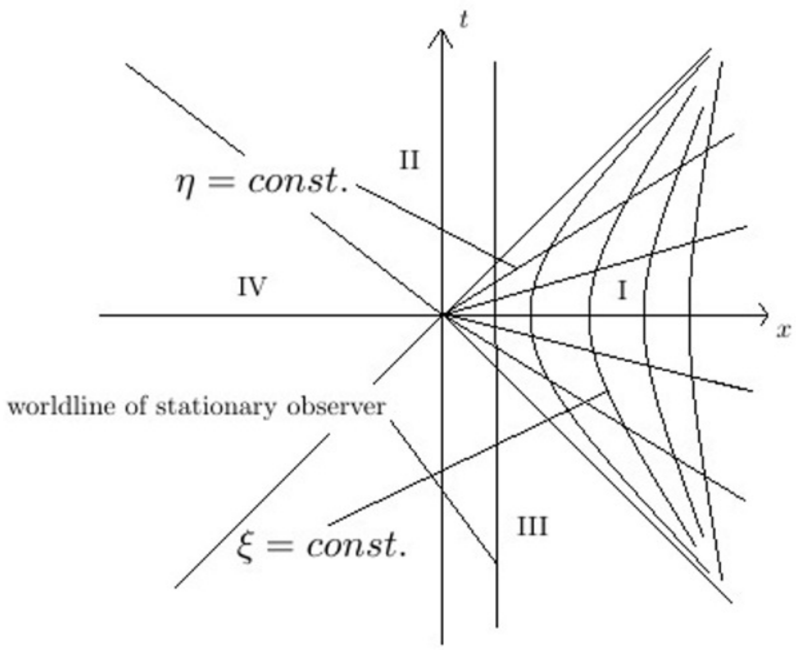

FIGURE 2.1: Rindler spacetime. Region I is the only region accessible for an eternally accelerating observer in the positive x-direction. Lines at 45 degrees demarcate separate regions.

are fundamentally different from black hole event horizons as the Rindler horizon is dependent on the observer's motion [12]). Due to the equivalence principle of general relativity these tools can apply to an observer in a constant gravitational field – we will elaborate on this in section 3.2 and build on it in our derivation of Hawking radiation.

The Schwarzschild black hole, although useful pedagogically, is not representative of the black holes thought to be found in our universe. Real black holes are predicted to have non-zero angular momentum and can also have electric charge, whereas the only property of the hole described by the Schwarzschild system is its mass. The metric used to describe a charged black hole is the Reissner-Nördstrom solution, this is formed by varying the Einstein-Maxwell action [12]

$$S = \frac{1}{16\pi} \int \sqrt{-g}\, (R - F_{\mu\nu}F^{\mu\nu})\, d^4x \tag{2.15}$$

to obtain the equations of motion

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 2\left(F_{\mu\alpha}F^{\alpha}_{\ \nu} - \frac{1}{4}g_{\mu\nu}F_{\beta\gamma}F^{\beta\gamma}\right), \tag{2.16}$$

$$\nabla_{\mu}F^{\mu\nu} = 0. \tag{2.17}$$

The spherically symmetric metric solution is [12]:

$$ds^2 = -\left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)dt^2 + \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \tag{2.18}$$

where Q is the charge of the black hole. This is the unique spherically symmetric solution to equations (2.16) and (2.17) [38]. Introducing the function:

$$\Delta = Q^2 - 2Mr + r^2 = (r - r_+)(r - r_-), \tag{2.19}$$

$$\text{where} \quad r_{\pm} = M \pm \sqrt{M^2 - Q^2}, \tag{2.20}$$

gives a form for the metric:

$$ds^2 = -\frac{\Delta}{r^2}dt^2 + \frac{r^2}{\Delta}dr^2 + r^2 d\theta^2 + r^2 sin^2\theta d\phi^2. \tag{2.21}$$

The units are *geometrised* in this subsection, so mass M and charge Q have the same units [12].

The cases where $Q > M$, $Q = M$, and $Q < M$ lead to very different physics and interesting reviews of these can be found in [12, 15, 38, 39]. How can a more general black hole be described?

For this we require the Kerr-Newman solution which includes both the charge *and* angular momentum of the hole. This solution describes a charged hole of mass M rotating through the polar angle ϕ with angular momentum per unit mass a [12, 39]:

$$\begin{aligned} ds^2 = -\left(\frac{\Delta - a^2 sin^2\theta}{\Sigma}\right)dt^2 + \frac{\Sigma}{\Delta}dr^2 - \frac{2a sin^2\theta}{\Sigma}\left(r^2 + a^2 - \Delta\right)dt d\phi + \Sigma d\theta^2 \\ + \left(\frac{\left(r^2 + a^2\right)^2 - \Delta a^2 sin^2\theta}{\Sigma}\right) sin^2\theta d\phi^2, \end{aligned} \tag{2.22}$$

where

$$\Sigma = r^2 + a^2 cos^2\theta, \tag{2.23}$$

$$\Delta = r^2 - 2Mr + Q^2 + a^2, \tag{2.24}$$

and the components of the vector potential are

$$A_t = \frac{Qr}{\Sigma}, \quad A_\phi = -\frac{Qarsin\theta}{\Sigma}, \quad A_r = A_\theta = 0. \tag{2.25}$$

When the charge Q is zero then we find the Kerr metric. The Kerr metric has the curious property that it describes a singularity in the shape of a ring centred at $r = 0$ with finite radius [12].

The Kerr/Kerr-Newman black hole descriptions are not spherically symmetric as they rotate around polar axes, this suggests that we cannot use an anaologue of Birkhoff's theorem, and so the metric shown above is not necessarily valid on the surface of the collapsing matter used to form the black hole [12]. At late times the spacetime around the rotating black hole "settles down" to a stationary state which exhibits time-translation symmetry or equivalenty has a timelike Killing vector field (introduced in the next section) [38].

2.2 Useful Mathematics for Black Holes

I have mainly followed the treatment of the mathematics of black holes given in *Carroll* [15] and *Poisson* [45]. One of the most powerful techniques used in the study of black holes and general relativity is differential geometry. Here we will look at some of the basics of differential geometry and introduce the technical apparatus required to understand some of the salient features of black holes.

A manifold is a topological space which locally looks Euclidean [46]. Let's define a vector field on the manifold $v^\mu(x)$, with integral curves $x^\mu(t)$ where $v^\mu = \frac{dx^\mu}{dt}$. Now if we have a tensor defined on the manifold at a certain point $T(p)$ (the arbitrary covariant and contravariant components of T are implicit) we would like to see how it changes with respect to the integral curves. A diffeomorphism maps a manifold to itself $\phi : M \to M$ and lets us push forward and pull back arbitrary tensors along integral curves; if the vector field on M is everywhere smooth and non-zero then the collection of its integral curves forms a congruence [15, 47].

$\phi_s(p)$ defines the point parameter distance s along from p; the parameter s indexes the distance along an integral curve (every point on the manifold is defined to be on a unique integral curve). The Lie derivative of the tensor T along the vector field $v^\mu(x)$ is given by [12]:

$$\mathcal{L}_v T = \lim_{t \to 0} \left(\frac{\phi_t^*[T(\phi_t(p))] - T(p)}{t} \right). \tag{2.26}$$

ϕ_t^* pulls back tensors. So the Lie derivative can be interpreted as comparing how the tensor at point p, $T(p)$, compares to the tensor at point $\phi_t(p)$ pulled back to point p along an integral curve – see Figure 2.2.

Let's apply the Lie derivative to the metric tensor $g_{\mu\nu}$ and utilise the equality [45]:

$$(\mathcal{L}_v g)_{\mu\nu} = \nabla_\mu v_\nu + \nabla_\nu v_\mu. \tag{2.27}$$

When the above equation equals zero this is known as Killing's equation and any vector v that satisfies it is a Killing vector. Killing vectors indicate the directions in which the metric is unaltered [48]. There is an easy way to find out whether a metric has a Killing vector: if the metric coefficients are independent of a coordinate, say for example time t, then $\frac{\partial}{\partial t}$ is a Killing vector. If the metric coefficients are independent of t then the metric is *stationary*, if they are independent of the polar angle ϕ then the solution is *axisymmetric* [49]. Spacetimes that are stationary and axisymmetric are significant in that they describe the equilibrium setting for rotating, axisymmetric bodies; and stationary, axisymmetric vacuum solutions to Einstein's field equations are very useful for describing the region around rotating black holes – cf. the Kerr metric [38].

Killing vectors also allow a conservation law to be formed for particles travelling on geodesics through a spacetime. For the Killing vector v^μ, a conserved charge Q is formed by multiplying the vector by the generalised momentum

$$p_\mu = \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \tag{2.28}$$

where $\mathcal{L}$ in equation (2.28) is the Lagrangian density of the system [12]:

$$Q = v^\mu p_\mu. \tag{2.29}$$

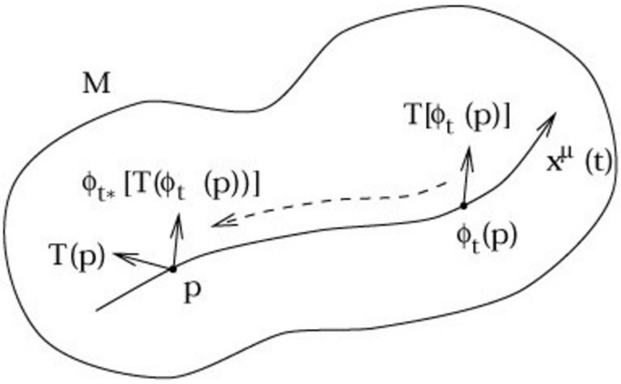

FIGURE 2.2: Pictorial view clarifying Lie derivative. Illustration from [50].

Hypersurfaces are also a prevalent feature of general relativity [45]. The hypersurfaces we will be looking at are submanifolds of one less dimension than the full four-dimensional spacetime. They can be either timelike, spacelike, or null, depending on the properties of their normal vectors. If one has a normal vector that is spacelike then the hypersurface is timelike and vice versa; a null hypersurface has null normal vector [12]. A very important feature of general relativity useful in the study of black holes is a *Killing horizon*. Given a Killing vector field χ^μ that is null along a null hypersurface Σ, then Σ is a Killing horizon of χ [45]. The link between Killing horizons and event horizons is given in Carroll [15]: "Every event horizon Σ in a stationary, asymptotically flat spacetime is a Killing horizon for some Killing vector field χ^μ". Killing horizons are useful in that they allow us to define a quantity known as *surface gravity*: along a Killing horizon, a Killing vector field χ^μ satisfies

$$\chi^\mu \nabla_\mu \chi^\alpha = -\kappa \chi^\alpha \tag{2.30}$$

where κ represents the surface gravity, which normally has a constant value over the horizon – except in certain special circumstances where it can change sign which will not be relevant for our discussion [15]. κ can be scaled by an arbitrarily chosen constant, however we can fix it at a certain value using boundary conditions. For example in an asymptotically flat, static spacetime the Killing vector $k = \frac{\partial}{\partial t}$ can be set to $k_\mu k^\mu = -1$ at infinity which fixes the value of κ. The applicability of the label "surface gravity" results from the following definition, also from *Carroll* [15]:

"In a static, asymptotically flat spacetime, the surface gravity is the acceleration of a static observer near the horizon, as measured by a static observer at infinity."

A static observer is defined to be one whose four-velocity u^μ obeys [15]

$$k^\mu = V u^\mu. \tag{2.31}$$

Multiplying the contravariant k^μ by its covariant counterpart k_μ, and remembering $u^2 = -1$, trivially gives a value for V (which is a function of spacetime coordinates) of $\sqrt{-k_\mu k^\mu}$. Due to our boundary conditions V is zero at Σ and tends to 1 as infinity is reached (since here $k_\mu k^\mu = -1$). V is generally called the *redshift factor* [38] for a reason we will now see.

Remember that conservation laws emerge from considerations of Killing vectors, from (2.29) we found that a Killing vector has an associated charge that is conserved along geodesics. In Schwarzschild spacetime, with metric given by (2.1), there is a conserved charge Q associated with Killing vector $k = \frac{\partial}{\partial t}$ [12] – in component form $k^\mu = (1, 0, 0, 0)$ – given by:

$$\begin{aligned} Q &= k^\mu p_\mu \\ &= k^\mu m g_{\mu\nu} \dot{x}^\nu \\ &= m g_{00} \dot{x}^0 \\ &= -m \left(1 - \frac{2M}{r}\right) \frac{dt}{d\tau} \end{aligned} \tag{2.32}$$

where m is the mass of the particle on the geodesic. To clarify the physical meaning of Q:

$$\begin{aligned} p_\mu &= (-E, \mathbf{p}) \\ &= m g_{\mu\nu} \frac{dx^\nu}{d\tau}, \end{aligned} \tag{2.33}$$

so the energy of the particle per its rest mass is (as measured by an inertial observer at infinity):

$$\begin{aligned} \frac{E}{m} &= -g_{00} \frac{dt}{d\tau} \\ &= \left(1 - \frac{2M}{r}\right) \frac{dt}{d\tau}, \end{aligned} \tag{2.34}$$

from here we trivially find $Q = -E$.

Back to our redshift factor V. The frequency of a photon as measured by an observer with four-velocity u^μ is $\omega = -p^\mu u_\mu$ [15], and from our proof above we have the photon's energy as $E = -p^\mu k_\mu$. The quotient $\frac{E}{V}$ therefore gives ω, revealing why V is called the redshift factor. Picture two static observers outside of a Schwarzschild black hole, observer 1 emits a photon of wavelength λ_1 which observer 2 measures as having wavelength redshifted to $\lambda_2 = \frac{V_2}{V_1}\lambda_1$.

The four-acceleration $a^\nu = u^\alpha \nabla_\alpha u^\nu$ can be expressed as $a_\nu = \nabla_\nu \ln V$ where

$$a = \frac{1}{V}\sqrt{\nabla_\nu V \nabla^\nu V}, \tag{2.35}$$

this goes to infinity at the horizon where the redshift factor becomes zero [15]. Schwarzschild spacetime is static (and therefore stationary) and asymptotically flat and so the Killing horizon associated with $\frac{\partial}{\partial t}$ coincides with its event horizon. Therefore we can see that a going to infinity at the Killing horizon is physically reasonable, since for an observer hovering *just above* the event horizon to stay static would require an incredibly high outward acceleration to ensure that they didn't fall in. At the horizon itself the acceleration would become infinite. The surface gravity at the horizon is taken to be

$$\kappa = Va, \tag{2.36}$$

the "redshifted" acceleration [38]. We can also see now that the event horizon is a surface of infinite redshift. Moving observer 1 to the horizon pushes its redshift factor V_1 to zero; infinitely stretching the outgoing signal.

Let's calculate some of the values for these quantities in Schwarzschild spacetime as they will be very useful for us later on when deriving black hole evaporation by Hawking radiation. The Schwarzschild spacetime has Killing vector $k^\mu = (1, 0, 0, 0)$ and a static four-velocity calculated by:

$$\begin{aligned} -1 &= u^\mu u^\nu g_{\mu\nu} \\ &= -\left(u^0\right)^2\left(1 - \frac{2M}{r}\right), \end{aligned} \tag{2.37}$$

using (2.1), leading to $u^0 = \frac{1}{\sqrt{1-\frac{2M}{r}}}$. From (2.31) V is obviously:

$$V = \sqrt{1 - \frac{2M}{r}} \tag{2.38}$$

Using $a_\nu = \nabla_\nu \ln V$ we find:

$$\begin{aligned} a_\nu &= \frac{1}{2}\nabla_\nu \left[\ln\left(1 - \frac{2M}{r}\right)\right] \\ &= \frac{M}{r^2\left(1 - \frac{2M}{r}\right)}\nabla_\nu r \\ &= \frac{M}{r^2\left(1 - \frac{2M}{r}\right)}\delta^r_\nu. \end{aligned} \tag{2.39}$$

Giving magnitude:

$$a = \frac{M}{r^2}\left(1 - \frac{2M}{r}\right)^{-\frac{1}{2}}. \tag{2.40}$$

Finally, the surface gravity at the horizon is equal to Va therefore using (2.38) gives for a Schwarzschild black hole [23]:

$$\kappa = \frac{1}{4M}. \tag{2.41}$$

2.3 Black Hole Thermodynamics

The deep connection between black holes and thermodynamics first came into view in the early 1970s [10]. There appeared to be a remarkable similarity between the "laws" that had been derived concerning the behaviour of black holes and the four laws of thermodynamics [20] – see Table 2.1. The "second law" for example posited that the entropy of a black hole was analogous to its *area*. Bekenstein [21] argued that black holes must have a physical entropy as otherwise throwing a highly entropic system into the black hole would allow for the entropy of the universe to be arbitrarily diminished. Therefore he put forward a generalised second law which stated that the total entropy of matter outside of black holes *plus* the entropy of the black holes themselves could never decrease. Initially it was believed that the similarity between the surface area of a black hole and its entropy was purely structural and that the black hole couldn't actually have an entropy as its temperature was zero [12]; however after Bekenstein's arguments and the discovery of Hawking radiation, a thermodynamic temperature was associated with the hole (which we shall derive in section 3.3), and the association of area and entropy

Law	Classical Thermodynamics	Black Holes
0th	The temperature T is constant all through a system in thermal equilibrium	The surface gravity κ remains constant over the event horizon of a stationary black hole
1st	$dE = TdS +$ work terms	$dM = \frac{1}{8\pi}\kappa dA + \Omega_H dJ$
2nd	The entropy S increases or stays the same in any process	The area A increases or stays the same in any process
3rd	$T = 0$ cannot be achieved in any physical process	$\kappa = 0$ cannot be achieved in any physical process

Table 2.1: Comparison of laws of thermodynamics and black hole mechanics [12, 20, 38].

was put onto more solid ground [24]. See *Carroll* [15] and *Jacobson* [23] for a much more detailed exposition of black hole thermodynamics.

Chapter 3

Hawking Radiation

3.1 Quantum Field Theory in Curved Spacetimes

Quantum field theory has over the past century proven to be the basis for arguably the most successful scientific theories ever produced. It forms the language of quantum electrodynamics, quantum chromodynamics and is a very powerful tool in the study and development of condensed matter systems [51]. The combination of quantum field theory and the general relativistic description of gravity has turned out to be extremely difficult, although significant progress has been made within the frameworks of supersymmetry, string theory, and other candidate theories [52–54]. In this section we will look at quantum field theory on fixed curved spacetimes. By fixing the background spacetime we significantly simplify the physics as we are in effect ignoring the "back reaction" [3] of the matter on the geometry. First we will very briefly look at the main features of a quantum field theory in flat spacetime, after which these features will be compared and contrasted with the features of the theory in a curved background. In this section I have mainly followed the treatments given in *Dowker* [12], *Carroll* [15], and *Wald* [55].

The simplest field to study, but containing all of the salient features relevant for our discussion, is the real massive scalar field ϕ. The Lagrangian density for ϕ is

$$\mathcal{L} = -\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2 \tag{3.1}$$

where the metric $g^{\mu\nu}$ here is Minkowskian and m is the field's mass [51]. Utilising the Euler-Lagrange equations yields the equation of motion:

$$\Box\phi = m^2\phi \tag{3.2}$$

known as the Klein-Gordon equation [51]. $\Box\phi$ denotes the d'Alembertian operator $\partial_\mu\partial^\mu$ acting on the field.

A plane wave solution to the Klein-Gordon equation is given by

$$\phi = \phi_0 e^{ip_\mu x^\mu}, \tag{3.3}$$

where $p^\mu = (\omega, \mathbf{k})$ with the time component representing angular frequency, the space components wave 3-vector, and satisfying the dispersion relation $p^2 + m^2 = 0$ (having set $\hbar = 1$). We wish to find the most general solution by forming an orthonormal basis into which any solution can be decomposed. This orthonormality is manifest by defining an inner product on the Klein-Gordon solution space. An inner product of two solutions over a $t =$ constant hypersurface is [15]

$$(\phi_1, \phi_2) = -i \int (\phi_1 \partial_t \phi_2^* - \phi_2^* \partial_t \phi_1)\, d^3x. \tag{3.4}$$

The complete orthonormal basis for solutions consists of all positive frequency plane wave modes and their complex conjugates $\{\psi_\mathbf{p}, \psi_\mathbf{p}^*\}$, i.e.

$$\phi = \int d^3p \left(a_\mathbf{p}\psi_\mathbf{p} + a_\mathbf{p}^\dagger \psi_\mathbf{p}^* \right), \tag{3.5}$$

where the coefficients and fields have been upgraded to operators under second quantisation (hats denoting operator status are implicit) [15]. $\psi_\mathbf{p}$ is positive frequency and its complex conjugate is negative frequency, they fulfill:

$$\partial_t \psi_\mathbf{p} = -i\omega\psi_\mathbf{p}, \tag{3.6}$$

$$\partial_t \psi_\mathbf{p}^* = i\omega\psi_\mathbf{p}^*, \tag{3.7}$$

$\omega > 0$ in both equations above even though (3.6) describes positive frequency modes and (3.7) describes negative frequency modes (this is because of the complex conjugate changing the sign of i in the plane wave solution) [15]. The coefficients are annihilation and creation operators satisfying

$$\begin{aligned}
[a_{\mathbf{p}}, a_{\mathbf{p}'}] &= 0, \\
[a^{\dagger}_{\mathbf{p}}, a^{\dagger}_{\mathbf{p}'}] &= 0, \\
[a_{\mathbf{p}}, a^{\dagger}_{\mathbf{p}'}] &= \delta^{(3)}\left(\mathbf{p} - \mathbf{p}'\right),
\end{aligned} \tag{3.8}$$

with the vacuum state defined by $a_{\mathbf{p}}|0\rangle = 0$ for all $\mathbf{p}$ [51]. These allow us to build up arbitrary quantum states of many particles (a Fock basis for the Hilbert space). $\omega > 0$ modes in an inertial frame can be decomposed into linear combinations of $\omega > 0$ modes in a Lorentz transformed frame, this means that all inertial observers will agree on whether they measure particles or whether they measure a vacuum state. A proof of this, following *Carroll* [15], shows how the crux of the issue lay in whether we can objectively call a mode's frequency positive or negative. Let's Lorentz boost a positive frequency mode by velocity $\mathbf{v} = \frac{d\mathbf{x}}{dt}$ to see how the frequency behaves in this new inertial frame (parameterized by $\mathbf{x}'$ and t') [38]:

$$t = \gamma\left(t' + \mathbf{v}.\mathbf{x}'\right), \tag{3.9}$$

$$\mathbf{x} = \gamma\left(\mathbf{x}' + \mathbf{v}t'\right). \tag{3.10}$$

In the boosted frame:

$$\partial_{t'}\psi_{\mathbf{p}} = \frac{\partial x^{\mu}}{\partial t'}\partial_{\mu}\psi_{\mathbf{p}} = -i\omega'\psi_{\mathbf{p}} \tag{3.11}$$

using the plane wave solution given in (3.3) and $\omega' = \gamma\left(\omega - \mathbf{v}.\mathbf{k}\right)$. This shows that a quantum particle defined on a flat background can be boosted to simply give the particle with boosted momentum in the new frame. The existence of particles or lack thereof (a vacuum) is therefore independent of our inertial frame in flat spacetimes and can be reliably defined. We will now see that this is not the case for field theory in curved spacetimes where the notions of "particle" and "vacuum" are shown to be observer-dependent.

When looking at field theories in curved backgrounds we must insist on the spacetime satisfying a condition known as *global hyperbolicity* [12] [55]. This requires that the spacetime under study, with specified geometry and metric, has a *Cauchy surface* Σ: a three-dimensional hypersurface that all past and future inextendible causal curves cross once.

Definition: *A causal curve has a tangent vector which is at no point space-like.*

Theorems 4.1.1. and 4.1.2 in *Wald* [55] state that a globally hyperbolic spacetime can be "foliated" by Cauchy surfaces such that each t = constant hypersurface is a Cauchy surface, and that quantum fields have solutions to their equations of motion defined throughout the spacetime from the initial data on Σ. This "foliation" is familiar from cosmology in the Robertson–Walker model [15]. Defining a real scalar field on Σ satisfying the Klein-Gordon equation given in (3.2) (where the d'Alembertian is now of a more complicated form than in the flat case due to the non-vanishing connection coefficients of the metric) and generalising the inner product from (3.4) gives:

$$(\phi_1, \phi_2) = -i \int (\phi_1 \nabla_\mu \phi_2^* - \phi_2^* \nabla_\mu \phi_1)\, dS^\mu, \tag{3.12}$$

where the integral is over the Cauchy surface Σ, dS is the infinitesimal induced volume element on the surface with normal n^μ, and $dS^\mu = dSn^\mu$ [15]. This inner product is independent of our choice of Σ.

Naively following our earlier flat spacetime procedure we would find solutions to the Klein-Gordon equation (now in a curved background) and decompose these into a basis of positive- and negative-frequency modes, with creation and annihilation operator coefficients. The problem with this method is that solutions of the Klein-Gordon equation which dissociate into space-dependent and time-dependent factors only exist when there is also a time-like Killing vector $k = \partial_t$ [38]. This dissociation is required in order to define positive- and negative-frequency modes in an invariant way [15]. We now see why unique positive- and negative-frequency solutions were found in flat spacetime: the Minkowski metric carries a time-like Killing vector $k = \partial_t$ and all k in different inertial frames are related by Lorentz boosts [15].

If a Cauchy surface exists in the spacetime then we can always decompose our solutions into orthonormal bases $\{f_i, f_i^*\}$, satisfying [12]:

$$\begin{aligned} (f_i, f_j) &= \delta_{ij}, \\ \left(f_i^*, f_j^*\right) &= -\delta_{ij}, \\ \left(f_i, f_j^*\right) &= 0. \end{aligned} \tag{3.13}$$

Choosing $\{f_i, f_i^*\}$ to be a complete set, we can expand ϕ as:

$$\phi = \sum_i \left(a_i f_i + a_i^\dagger f_i^* \right), \tag{3.14}$$

the coefficients satisfy:

$$\begin{aligned} [a_i, a_j] &= 0, \\ [a_i^\dagger, a_j^\dagger] &= 0, \\ [a_i, a_j^\dagger] &= \delta_{ij}. \end{aligned} \tag{3.15}$$

As before we define a_i to be an annihilation operator and $a_i^\dagger$ to be a creation operator, therefore the f-vacuum is defined by:

$$a_i|0_f\rangle = 0 \quad \forall i, \tag{3.16}$$

and we can build up a number of excitations in state i as usual:

$$|n_i\rangle = \frac{1}{\sqrt{n_i!}} \left(a_i^\dagger \right)^{n_i} |0_f\rangle. \tag{3.17}$$

The number operator for f-excitations is as expected:

$$n_{fi} = a_i^\dagger a_i. \tag{3.18}$$

Unlike in the flat case, the basis $\{f_i, f_i^*\}$ is highly nonunique and so the concepts of vacuum and number operator will be dependent on our choice of basis. Let's look at a different basis $\{g_i, g_i^*\}$, permitting the expansion:

$$\phi = \sum_i \left(b_i g_i + b_i^\dagger g_i^* \right), \tag{3.19}$$

the new annihilation b_i and creation $b_i^\dagger$ operators satisfy the same commutation relations

$$\begin{aligned} [b_i, b_j] &= 0, \\ [b_i^\dagger, b_j^\dagger] &= 0, \\ [b_i, b_j^\dagger] &= \delta_{ij}. \end{aligned} \tag{3.20}$$

With vacuum state defined by

$$b_i|0_g\rangle = 0 \quad \forall i, \tag{3.21}$$

and number operator $n_{gi} = b_i^\dagger b_i$.

To see how different observers measure the same physical phenomena in curved spacetimes, let's expand each mode in terms of the other one [15]:

$$g_i = \sum_j \left(\alpha_{ij} f_j + \beta_{ij} f_j^*\right), \tag{3.22}$$

$$f_i = \sum_j \left(\alpha_{ji}^* g_j - \beta_{ji} g_j^*\right). \tag{3.23}$$

These relations between different bases are examples of what is known as *Bogolubov transformations* [56]; α_{ij} and β_{ij} are Bogolubov coefficients, they satisfy $\alpha_{ij} = (g_i, f_i)$ and $\beta_{ij} = -\left(g_i, f_j^*\right)$, as well as [15]:

$$\begin{aligned}\sum_j \left(\alpha_{ik}\alpha_{jk}^* - \beta_{ik}\beta_{jk}^*\right) &= \delta_{ij},\\ \sum_j \left(\alpha_{ik}\beta_{jk} - \beta_{ik}\alpha_{jk}\right) &= 0.\end{aligned} \tag{3.24}$$

The Bogolubov transformation also allows us to express operators in one mode in terms of operators in another mode [15]:

$$\begin{aligned}a_i &= \sum_j \left(\alpha_{ji} b_j + \beta_{ji}^* b_j^\dagger\right),\\ b_i &= \sum_j \left(\alpha_{ij}^* a_j - \beta_{ij}^* a_j^\dagger\right).\end{aligned} \tag{3.25}$$

Now we are ready to see what an observer using one set of modes sees in a vacuum defined in another set of modes. Setting our quantum state to be in the f-vacuum, we find the expectation value of the g-mode number operator to be:

$$
\begin{aligned}
\langle 0_f | n_{gi} | 0_f \rangle &= \langle 0_f | b_i^\dagger b_i | 0_f \rangle \\
&= \left\langle 0_f \middle| \sum_{jk} \left(\alpha_{ij} a_j^\dagger - \beta_{ij} a_j \right) \left(\alpha_{ik}^* a_k - \beta_{ik}^* a_k^\dagger \right) \middle| 0_f \right\rangle \\
&= \sum_{jk} \left(-\beta_{ij} \right) \left(-\beta_{ik}^* \right) \langle 0_f | a_j a_k^\dagger | 0_f \rangle \\
&= \sum_{jk} \beta_{ij} \beta_{ik}^* \langle 0_f | \left(a_k^\dagger a_j + \delta_{jk} \right) | 0_f \rangle \\
&= \sum_{jk} \beta_{ij} \beta_{ik}^* \delta_{jk} \langle 0_f | 0_f \rangle \\
&= \sum_{j} \beta_{ij} \beta_{ij}^*
\end{aligned}
\tag{3.26}
$$

using equations (3.15) and (3.25). In general this expectation value will be non-zero and so the question of the existence of particles or the presence of a vacuum will depend ultimately on the observer. In quantum field theory, it is the *fields* that are the fundamental physical objects, the picture of particles are contingent and in some cases not a sharply defined concept [15].

A very simple but useful picture of how particles can be *created* by the changing curvature of a spacetime comes from looking at a wavefunction's response to a sudden change in potential [5]. Quantum field theory dictates that each mode of a field acts as a harmonic oscillator [51]. This mode has a certain frequency ω_1, however once the spacetime in which it dwells has changed (for example due to the influence of some mass) the mode's frequency can alter, giving a different value ω_2. Using the well-known adiabatic theorem of quantum mechanics we know that changing a potential "slowly" [57] with repect to the inverse of the frequencies ω_1 and ω_2, which we take to be of the same order [5], will allow the mode to adapt in such a way that it stays as the (say) vacuum state throughout the evolution of the potential. If instead the potential changes suddenly, the wavefunction has not had enough time to evolve appropriately and so it will change from a vacuum state to an excited state.

The ω_1 vacuum can be expanded in terms of the excited states of ω_2 as

$$
|0\rangle_{\omega_1} = C|0\rangle_{\omega_2} + C_1|1\rangle_{\omega_2} + C_2|2\rangle_{\omega_2} + C_3|3\rangle_{\omega_2} + ... \tag{3.27}
$$

but the wavefunctions are symmetric and so we can throw out the odd-numbered harmonic excitations in our expansion [5], leaving:

$$|0\rangle_{\omega_1} = C|0\rangle_{\omega_2} + C_2|2\rangle_{\omega_2} + ... \tag{3.28}$$

So a fast enough change in spacetime curvature will create pairs of excitations.

3.2 The Unruh Effect

In this section we'll use the notions developed previously of the relativity of particles and the vacuum, and build on the earlier exposition of Rindler spactime, to derive the result that an accelerating observer in a Minkowski vacuum will detect thermal radiation, known as *the Unruh effect*. This phenomenon will then be shown to imply the Hawking effect whereby a black hole emits thermal radiation – the source of the information paradox.

In section 2.1 we looked at Rindler space, the description of a constantly accelerating observer in flat spacetime. Now let's look into its structure in a bit more detail before quantising and expanding a scalar field with respect to the Rindler observers. In Figure 2.1 we saw that the Minkowski spacetime is divided into four regions by the Rindler coordinates with region I referred to as Rindler space. There is a strong parallel between Rindler observers in region I and static observers in the $r > GM$ section of Schwarzschild spacetime. Here again is the Rindler metric:

$$ds^2 = e^{2a\xi}\left(-d\eta^2 + d\xi^2\right), \tag{3.29}$$

none of the metric coefficients depend on η therefore there exists a Killing vector ∂_η, which in terms of Minkowski coordinates becomes:

$$\begin{aligned}\partial_\eta &= \frac{\partial t}{\partial \eta}\frac{\partial}{\partial t} + \frac{\partial x}{\partial \eta}\frac{\partial}{\partial x} \\ &= a\left(x\frac{\partial}{\partial t} + t\frac{\partial}{\partial x}\right)\end{aligned} \tag{3.30}$$

using equation (2.12). This is the Killing field for a Lorentz boost in the positive x-direction [15]. In regions I and IV of Figure 2.1 ∂_η is a time-like Killing vector whereas in regions II and III it is space-like. Therefore the lines $x = \pm t$ are both Killing horizons

[15, 43]. The Rindler coordinates in region IV can be defined as:

$$\begin{aligned} t &= -\frac{1}{a} e^{a\xi} \sinh(a\eta), \\ x &= -\frac{1}{a} e^{a\xi} \cosh(a\eta). \end{aligned} \tag{3.31}$$

where $x < |t|$. As in the last section, a time-like Killing vector is required to define positive- and negative-frequency modes in an invariant way, therefore here we can use ∂_η. Let's consider a massless 1+1 dimensional scalar field ϕ obeying the usual Klein-Gordon equation, expressed in Rindler coordinates:

$$\Box\phi = e^{-2a\xi} \left(\frac{\partial^2}{\partial\eta^2} + \frac{\partial^2}{\partial\xi^2} \right) \phi = 0, \tag{3.32}$$

solved by plane wave mode:

$$g_k = \frac{1}{\sqrt{4\pi\omega}} e^{-i\omega\eta + ik\xi}, \tag{3.33}$$

where $\omega = |k|$ [15]. From equations (2.13) it is noted that η is proportional to the Rindler observer's proper time and so our definition of a positive-frequency mode should be with respect to η:

$$\frac{\partial}{\partial\eta} g_k = -i\omega g_k \tag{3.34}$$

as required cf. (3.6). One caveat is that the associated Killing vector defining the positive-frequency modes g_k must be future-directed [15]. $\frac{\partial}{\partial\eta}$ points in opposite directions in regions I and IV, so the Killing vector required to define positive frequency modes in region IV must be $-\frac{\partial}{\partial\eta}$ (for clarity: $\frac{\partial}{\partial\eta}$ is future-directed in region I, $-\frac{\partial}{\partial\eta}$ is future-directed in region IV). To cope with this we define two sets of modes, each one non-zero in a certain region [15]:

$$g_k^{(1)} = \frac{1}{\sqrt{4\pi\omega}} e^{-i\omega\eta + ik\xi}. \tag{3.35}$$

Note that equation (3.35) is only valid in region I, in region IV $g_k^{(1)} = 0$. Also [15]

$$g_k^{(2)} = \frac{1}{\sqrt{4\pi\omega}} e^{i\omega\eta + ik\xi} \tag{3.36}$$

where (3.36) is valid only in region IV, $g_k^{(2)} = 0$ in region I. Taking these two sets of modes together we see that both regions I and IV are covered; the Rindler metric (2.14) is valid throughout these two regions [12, 15]. Both sets of modes can now be seen to be positive-frequency with respect to future-directed Killing vectors:

$$\begin{aligned} \frac{\partial}{\partial \eta} g_k^{(1)} &= -i\omega g_k^{(1)}, \\ -\frac{\partial}{\partial \eta} g_k^{(2)} &= -i\omega g_k^{(2)}. \end{aligned} \tag{3.37}$$

Where $\omega > 0$. Compare to equations (3.6), (3.7) to see that both expressions in (3.37) do indeed correspond to positive-frequency modes.

These modes and their complex conjugates form a complete set throughout the spacetime under study; expanding a scalar field in these modes gives [15]:

$$\phi = \int dk \left(b_k^{(1)} g_k^{(1)} + b_k^{(1)\dagger} g_k^{(1)*} + b_k^{(2)} g_k^{(2)} + b_k^{(2)\dagger} g_k^{(2)*} \right). \tag{3.38}$$

The mode coefficients are the annihilation and creation operators associated with their regions of validity. The modes themselves satisfy inner product identities analogous to (3.13) given by:

$$\begin{aligned} \left(g_{k_1}^{(1)}, g_{k_2}^{(1)} \right) &= \delta\left(k_1 - k_2\right) & \left(g_{k_1}^{(1)*}, g_{k_2}^{(1)*} \right) &= -\delta\left(k_1 - k_2\right) \\ \left(g_{k_1}^{(2)}, g_{k_2}^{(2)} \right) &= \delta\left(k_1 - k_2\right) & \left(g_{k_1}^{(2)*}, g_{k_2}^{(2)*} \right) &= -\delta\left(k_1 - k_2\right) \\ \left(g_{k_1}^{(1)}, g_{k_2}^{(2)} \right) &= 0 & \left(g_{k_1}^{(1)*}, g_{k_2}^{(2)*} \right) &= 0. \end{aligned} \tag{3.39}$$

Comparing our earlier field expansion in Minkowski modes (3.5) (at the time we defined it in terms of 3+1 dimensions, now we will treat it in 1+1 dimensions) with our new expansion in terms of Rindler modes (3.38) we can note that the vacuums do not coincide. For example the vacuum as described by an observer in Minkowski spacetime $a_k|0_M\rangle = 0$ will be seen to be bubbling with particles by a Rindler observer who, in turn, will describe a vacuum by $b_k^{(1)}|0_R\rangle = b_k^{(2)}|0_R\rangle = 0$ that the Minkowskian observer will see as anything but empty. This is as a result of the lack of a purely positive-frequency Minkowski mode basis with which to expand a Rindler mode; for example the annihilation operator defining the Rindler vacuum can only be decomposed into a combination of Minkowski annihilation *and* creation operators [15].

Now let's investigate the properties of the particles that a Rindler observer would see as they accelerated through the Minkowski vacuum. Rather than finding the Minkowski vacuum expectation value of the Rindler number operator using Bogolubov coefficients as in section 3.1, we'll utilise a shorter derivation based on the analytic continuation of the Rindler modes to the whole spacetime and subsequently expressing this result in terms of the more restricted Rindler modes valid only in regions I and IV. The derivation presented here follows that of *Dowker* [12] and *Carroll* [15].

Using the coordinate definitions given in (2.12) and (3.31) and some light algebra we find [12]:

$$e^{-a(\eta-\xi)} = \begin{cases} a\,(-t+x) & \text{I} \\ a\,(t-x) & \text{IV} \end{cases} \tag{3.40}$$

$$e^{a(\eta+\xi)} = \begin{cases} a\,(t+x) & \text{I} \\ a\,(-t-x) & \text{IV} \end{cases} \tag{3.41}$$

From the expression for mode $g_k^{(1)}$ in (3.35) we derive:

$$\begin{aligned} \sqrt{4\pi\omega}\, g_k^{(1)} &= e^{-i\omega\eta+ik\xi} \\ &= e^{-i\omega(\eta-\xi)} \\ &= a^{\frac{i\omega}{a}}\,(-t+x)^{\frac{i\omega}{a}}\,, \end{aligned} \tag{3.42}$$

where we've used the fact that $\omega = k$ which is valid because in our plane wave expression for g_k (3.33) we had $\omega = |k|$, and here $k > 0$. The analytic continuation of (3.42) is simply attained by applying it for any values of the Minkowski coordinates (t, x) [15]. As we stated above, our method here is to analytically continue the Rindler modes to the whole spacetime and then to express these in terms of the original Rindler modes in regions I and IV; however the continuation we have just found involves $g_k^{(1)}$ which vanishes in region IV, so we must consider the other modes $g_k^{(2)}$ too in order to cover both regions. Repeating our procedure above we obtain:

$$\begin{aligned} \sqrt{4\pi\omega}\, g_k^{(2)} &= e^{i\omega\eta+ik\xi} \\ &= e^{i\omega(\eta+\xi)} \\ &= a^{-\frac{i\omega}{a}}\,(-t-x)^{-\frac{i\omega}{a}}\,. \end{aligned} \tag{3.43}$$

We want the right sides of (3.42) and (3.43) to match up and so, making us of $e^{-i\pi} = -1$, we transform (3.43) thusly:

$$\begin{aligned} \sqrt{4\pi\omega}g_{-k}^{(2)*} &= e^{-i\omega\eta+ik\xi} \\ &= e^{-i\omega(\eta-\xi)} \\ &= a^{\frac{i\omega}{a}}\left(t-x\right)^{\frac{i\omega}{a}} \\ &= a^{\frac{i\omega}{a}}\left[e^{-i\pi}\left(-t+x\right)\right]^{\frac{i\omega}{a}} \\ &= a^{\frac{i\omega}{a}}e^{\frac{\pi\omega}{a}}\left(-t+x\right)^{\frac{i\omega}{a}}, \end{aligned} \tag{3.44}$$

where * denotes complex conjugation. Now we can express a combination of modes as [15]:

$$\sqrt{4\pi\omega}\left(g_k^{(1)} + e^{-\frac{\pi\omega}{a}}g_{-k}^{(2)*}\right) = a^{\frac{i\omega}{a}}\left(-t+x\right)^{\frac{i\omega}{a}}. \tag{3.45}$$

The mode given above covers regions I and IV as we wanted however (3.45) is not normalised, as an ansatz [15] let's assume the normalised analytic continuations of $g_k^{(1)}$ and $g_k^{(2)}$ are given by:

$$\begin{aligned} h_k^{(1)} &= \frac{1}{\sqrt{2\sinh\left(\frac{\pi\omega}{a}\right)}}\left(e^{\frac{\pi\omega}{2a}}g_k^{(1)} + e^{-\frac{\pi\omega}{2a}}g_{-k}^{(2)*}\right), \\ h_k^{(2)} &= \frac{1}{\sqrt{2\sinh\left(\frac{\pi\omega}{a}\right)}}\left(e^{\frac{\pi\omega}{2a}}g_k^{(2)} + e^{-\frac{\pi\omega}{2a}}g_{-k}^{(1)*}\right). \end{aligned} \tag{3.46}$$

This normalisation can be verified by checking the mode inner products identical in form to (3.39): i.e. $\left(h_{k_1}^{(1)}, h_{k_2}^{(1)}\right) = \delta\left(k_1 - k_2\right)$ etc. Expanding a scalar field ϕ in our normalised modes gives:

$$\phi = \int dk\left(c_k^{(1)}h_k^{(1)} + c_k^{(1)\dagger}h_k^{(1)*} + c_k^{(2)}h_k^{(2)} + c_k^{(2)\dagger}h_k^{(2)*}\right). \tag{3.47}$$

In complete analogy to our discussion earlier of Bogolubov transformations in section 3.1 – cf. equations (3.22), (3.23), (3.25) – we can express our original Rindler operators b_k in terms of our new extended operators c_k as:

$$\begin{aligned} b_k^{(1)} &= \frac{1}{\sqrt{2\sinh\left(\frac{\pi\omega}{a}\right)}}\left(e^{\frac{\pi\omega}{2a}}c_k^{(1)} + e^{-\frac{\pi\omega}{2a}}c_{-k}^{(2)\dagger}\right), \\ b_k^{(2)} &= \frac{1}{\sqrt{2\sinh\left(\frac{\pi\omega}{a}\right)}}\left(e^{\frac{\pi\omega}{2a}}c_k^{(2)} + e^{-\frac{\pi\omega}{2a}}c_{-k}^{(1)\dagger}\right). \end{aligned} \tag{3.48}$$

An accelerating Rindler observer in region I will form a number operator $n_{\mathrm{R}}^{(1)}(k) = b_k^{(1)\dagger} b_k^{(1)}$, which we can now express in terms of c_k operators. The new positive-frequency modes h_k can be decomposed into positive-frequency Minkowski modes [12] and so their vacuum states are identified. This allows us to finally obtain the expectation value for a region I Rindler observer in the Minkowski vacuum:

$$\begin{aligned} \langle 0_{\mathrm{M}} | n_{\mathrm{R}}^{(1)}(k) | 0_{\mathrm{M}} \rangle &= \langle 0_{\mathrm{M}} | b_k^{(1)\dagger} b_k^{(1)} | 0_{\mathrm{M}} \rangle \\ &= \frac{1}{2\mathrm{sinh}\left(\frac{\pi\omega}{a}\right)} \langle 0_{\mathrm{M}} | e^{-\frac{\pi\omega}{a}} c_{-k}^{(2)} c_{-k}^{(2)\dagger} | 0_{\mathrm{M}} \rangle \\ &= \frac{e^{-\frac{\pi\omega}{a}}}{2\mathrm{sinh}\left(\frac{\pi\omega}{a}\right)} \delta(0) \\ &= \frac{1}{e^{\frac{2\pi\omega}{a}} - 1} \delta(0), \end{aligned} \tag{3.49}$$

the $\delta(0)$ factor is an artifact of our choice of basis and can be transformed away (has no physical meaning) [15]. The expectation value (3.49) is a Planckian distribution at temperature [10]

$$T = \frac{a}{2\pi} \tag{3.50}$$

where $c = \hbar = 1$ and a is the magnitude of the acceleration. So a constantly accelerating observer through the Minkowski vacuum experiences a thermal bath of particles.

3.3 Black Hole Evaporation

The core of the information paradox lies in Hawking radiation. If black holes behaved classically – i.e. there was no evaporation mechanism – then the information that was consumed by the hole would exist just out of reach, but the information would not necessarily be destroyed. However a quantum mechanical treatment of black holes suggests that they eventually (over time scales many orders of magnitude larger than the age of the universe) vanish [24], leaving behind thermal radiation, seemingly erasing the information that fell in.

A heuristic picture of Hawking radiation involves a pair of photons emerging from the vacuum, one of which has energy $-E$, due to a fluctuation near the event horizon of a black hole. If this process occurs far away from any black holes the photons would annihilate in a time of order $\sim \frac{\hbar}{E}$. However if the pair creation happens close to the horizon then the photon of energy $-E$ may fall into the black hole whereas the other photon can travel to infinity, where an observer may observe it as Hawking radiation [10].

Rather than following Hawking's original derivation of black hole radiation [24] we'll see how our derivation of Unruh radiation leads to the existence of Hawking radiation due to the equivalence principle [15, 23].

Let's assume there are two static observers outside of a black hole in Schwarzschild spacetime: one at r_1, and another one further out at r_2 which we will push to infinity. In the following calculations we will only be using Schwarzschild spacetime as it contains all of the pertinent properties, we must also set some physical scales with which to anchor our derivation. The natural acceleration scale is $\frac{1}{2M}$ the inverse of the Schwarzschild radius, with the radius itself setting the natural time, length, and curvature scales near the horizon [15]. For a static observer just outside the event horizon at r_1, its acceleration is huge with respect to $\frac{1}{2M}$:

$$a_1 \gg \frac{1}{2M}, \tag{3.51}$$

and from the inverse $\frac{1}{a_1} \ll 2M$ we see that the observer at r_1 experiences an almost flat spacetime environment. From the equivalence principle we know that someone falling through the event horizon and into a black hole feels nothing strange (ignoring tidal forces) and so they experience approximately a Minkowski vacuum. Now we begin to see how the Unruh radiation manifests itself in this situation: a freely-falling observer sees a vacuum however a static observer at r_1 is accelerating outwardly with magnitude a_1 in order to stay at the same radial distance and so experiences Unruh radiation at temperature $\frac{a_1}{2\pi}$, cf. (3.50).

Another observer stationed at r_2 which we will assume is infinitely far away doesn't necessarily dwell in a flat region of spacetime like the observer stationed just outside the event horizon. Therefore they shouldn't experience Unruh radiation as the derivation of that was predicated on acceleration through *flat* spacetime. However the Unruh radiation detected by the observer at r_1 will continue to propagate outwards towards the observer at infinity, although in a redshifted form [15].

The amount of redshift will be dictated by the redshift factor V introduced earlier in section 2.2:

$$T_2 = \frac{V_1}{V_2} T_1, \tag{3.52}$$

where the V's and T's refer to the redshift factors and thermal radiation temperatures experienced by observers at r_1 and r_2. We know that $T_1 = \frac{a_1}{2\pi}$ and that at infinity the redshift factor tends to unity, therefore the temperature of the radiation experienced by the faraway observer – using (2.36) and letting observer 1's position r_1 be infinitesimally close to the horizon where $V_1 a_1$ tends to κ – is

$$T = \frac{\kappa}{2\pi}. \tag{3.53}$$

This is the temperature of the Hawking radiation emitted by the black hole.

Chapter 4

The Structure of Information

4.1 What is Information?

When one first hears the word "information" heard in a scientific context it can be confusing as to what exactly is meant by the term. It sounds like a very subjective concept which would be hard to quantify. To understand information we must first see how it relates to entropy.

Entropy is one of the most fundamental and important concepts in physics; useful not only in thermodynamics but in information theory, quantum theory, and as a basis for an *arrow of time* [58]. It is a measure of the *disorder* inherent in a system. It was first discovered (or invented) as a thermodynamical quantity related to energy and temperature. After the development of statistical mechanics the idea of entropy was detached and expanded somewhat and understood more in terms of the internal configurations of systems, along with their respective probabilities of occurance [58]. The second law of thermodynamics states that the entropy of the universe can never decrease and was said by Arthur Eddington to hold "the supreme position among the laws of Nature" [59]. So how does entropy relate to information?

The entropy of a thermodynamic system is:

$$-k_{\mathrm{B}}\sum_{i}p_{i}\ln p_{i} \tag{4.1}$$

where the sum is over all possible microstates corresponding to a given macrostate of the system, each microstate has a probability p_i of occuring. If we assume that each microstate has equal probability of occurrence $p_i = \frac{1}{N}$, N being the number of microstates corresponding to a given macrostate, then we recover the famous equation due to Boltzmann:

$$S = k_{\mathrm{B}} \mathrm{ln} N \tag{4.2}$$

valid at thermodynamic equilibrium in the micro-canonical ensemble [60].

The notion of information as formulated by Hartley and Shannon is a *generalisation* of this notion of entropy. Ralph Hartley (1888-1970), Claude Shannon (1916-2001), and others invented the modern theory of information [61, 62] originally as a way to speed up, compress, and process the signals used in communications. It has since been generalised to incorporate quantum theory and has been fundamental in the development of quantum computing and the study of entanglement [63].

Ralph Hartey was the first person to quantify the information contained in a message source (an ensemble of messages) in 1928 [58]. He did so using two parameters: the number of characters in the message n, and the number of equiprobable symbols that each character may adopt s. So for example, for the bit string 1011001110, $n = 10$ and $s = 2$. Hartley developed the concept of the information of a message source using the following assumptions: the information must be a function of both s and n, must be proportional to the message length n, and should monotonically increase with the number of equiprobable messages s^n. Representing the information by the symbol H and using the previous assumptions:

$$H = nf(s) = g(s^n), \tag{4.3}$$

where both $f(s)$ and $g(s^n)$ are monotonically increasing on their respective domains. The only differentiable solution is:

$$f(s) = \mathrm{cln}s \tag{4.4}$$

where c is a positive constant. (More accurately H represents the *missing information*, it denotes the missing information needed to select a message from the message source.) Plugging the solution into equation (4.3) it is easy to see that $H = \mathrm{cln}s^n$ – notice this is redolent of equation (4.2). We can see that the number of messages s^n in the message source is analogous to the number of microstates N of the thermodynamic system in equation (4.2), and that $\frac{S}{k_\text{B}} = \ln N$ is the missing information required to specify the microstate corresponding to a given macrostate.

Claude Shannon generalised Hartley's concept of information to include messages made up of characters that have unequal probabilities of occuring, this gives a form for the Shannon information [58]:

$$H = -\mathrm{c}\sum_{i=1}^{s} p_i \ln p_i \tag{4.5}$$

which is analogous to equation (4.1).

All of this seems to strongly suggest that missing or hidden information is in some way equivalent to entropy. For example a sealed box of hot gas could be in many different microstates with respect to its bulk thermodynamics quantities, so we know very little (hold limited information) about the internal configurations of the system and therefore its entropy is large. On the other hand if we were studying a perfectly ordered crystal at absolute zero temperature with a non-degenerate ground state (cf. the third law of thermodynamics) we would have complete knowledge of its internal configuration and so the entropy would be zero and there would be no hidden information [60]. Before we see how all of this relates to black holes we must look at the generalisation of entropy (missing information) to quantum theory: the von Neumann entropy.

In the 1930s [64] John von Neumann reformulated quantum mechanics by replacing the idea of a wave function – defining the state of a quantum system – by an object called a *density matrix*, describing a statistical ensemble of different quantum states. This allowed the apparatus of classical statistical mechanics to be extended to quantum theory. The density matrix is used to describe generic quantum systems that are either *pure states* or *mixed states*. If we have complete knowledge of a quantum state then we know that the probability of it being in that state is unity [65]. In reality it is more common to have incomplete knowledge of the quantum state(s) and/or there may be an ensemble of different states [66].

The density matrix ρ, defined by

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{4.6}$$

tells us whether the states are pure or mixed. For a generic mixed state, the wave function ψ_i occurs with probability p_i, and the expectation value of an observable described by operator O is:

$$\langle O \rangle = \sum_i p_i \langle \psi_i | O | \psi_i \rangle. \tag{4.7}$$

Expanding the states in a basis α_i [67]:

$$|\psi_i\rangle = \sum_j |\alpha_j\rangle\langle\alpha_j|\psi_i\rangle \tag{4.8}$$

$$\langle\psi_i| = \sum_k \langle\psi_i|\alpha_k\rangle\langle\alpha_k| \tag{4.9}$$

and substituting the expansions into equation (4.7) gives

$$\langle O \rangle = \sum_{j,k} \left(\sum_i p_i \langle\alpha_j|\psi_i\rangle\langle\psi_i|\alpha_k\rangle \right) \langle\alpha_k|O|\alpha_j\rangle \tag{4.10}$$

$$= \sum_{j,k} \langle\alpha_j|\rho|\alpha_k\rangle\langle\alpha_k|O|\alpha_j\rangle \tag{4.11}$$

$$= tr\left(\rho O\right). \tag{4.12}$$

Density matrices have the following properties: their trace is always equal to one (because the probabilities add up to one), they are Hermitian, and their eigenvalues are equal to or greater than zero. When ρ is diagonalised, the eigenvalues ρ_i represent the probabilities that the quantum system is in state i. As mentioned above, if a system is in a pure state then we possess complete knowledge of it, therefore the density matrix will have one non-zero eigenvalue (due to the trace property it must be equal to one) as there is a unity probability of it being in this eigenstate [3]. If the density matrix has more than one non-zero eigenvalue then the state is mixed.

The *von Neumann entropy* is an extension of equation (4.1) to quantum systems and is given by:

$$S = -\mathrm{tr}\rho\ln\rho = -\sum_i \rho_i \ln\rho_i \tag{4.13}$$

where tr gives the trace. S is a quantitative measure of how mixed a quantum state is [66]. A pure state has zero entropy whereas mixed states have entropy greater than zero.

The von Neumann entropy is also very useful as it quantifies the degree of entanglement between quantum subsystems [68]. To illustrate this let's look at a pure state Ψ with density matrix $\rho_{total} = |\Psi\rangle\langle\Psi|$. As this is a pure state it must have von Neumann entropy equal to zero. However let's now split the system into two parts A and B; the Hilbert space containing the states is also split into two: $\mathcal{H}_{total} = \mathcal{H}_A \otimes \mathcal{H}_B$. If an observer can only access subsystem A then he will measure the density matrix to be [68]:

$$\rho_A = \mathrm{tr}_\mathrm{B}\rho_{total}, \tag{4.14}$$

tr_B signifies that the trace is over $\mathcal{H}_B$. The von Neumann entropy of ρ_A defines the entanglement entropy of subsystem A [68]:

$$S_A = -\mathrm{tr}_A\rho_A\ln\rho_A. \tag{4.15}$$

Entanglement entropy illustrates a very important point about how entropy behaves differently in quantum systems. The entropy S of a system as defined by Shannon cannot be lower than the entropy of any of its constituent parts, whereas the von Neumann entropy for an entangled system – for example a Bell pair – has total entropy equal to zero but subsystem entropies greater than zero [69, 70]. We can know everything about a system yet know nothing about its parts – a very quantum phenomenon. In the next section we will see how the concepts introduced here relate to the problem (or illusion) of information loss in black hole evaporation.

4.2 Throwing a Bit of Information into a Black Hole

To get to grips with the information content of a black hole let's look at how the *area* of its event horizon changes when a single bit of information is dropped in. How does one drop a single bit of information into a black hole?

If we tried to encode a single bit in terms of say a 1 or 0 written on a piece of paper and then threw this into a black hole this wouldn't suffice because the mark on the paper would be formed from many atoms, this would still be the case even if the paper it was written on had nanoscale dimensions. The solution is to use an elementary particle, say a photon; however as *Susskind* [28] points out a single photon could still encode more than one bit of information if we had knowledge of the location of the point of entry into the black hole. Using a photon of wavelength of order the Schwarzschild radius of the hole results in the scenario that a photon has entered the black hole at some location on the event horizon, but it is not known where – this encodes one bit of information [28].

Using the basic Planck–Einstein relation $E = hf$ and Einstein's $E = mc^2$ [71] we find that the energy of a photon of wavelength Schwarzschild radius R_S is

$$E = \frac{hc}{R_S}, \tag{4.16}$$

and that after the photon has entered the black hole the hole's energy has increased by E and its mass has increased by $\frac{E}{c^2} = \frac{h}{cR_S}$. Therefore the increase in the event horizon's radius is (temporarily restoring units)

$$\Delta R_S = \frac{2Gh}{c^3 R_S}. \tag{4.17}$$

From the event horizon area formula $A = 4\pi R_S^2$ we find that the horizon area increases by

$$\begin{aligned}
\Delta A &= 4\pi \left(R_S + \Delta R_S\right)^2 - 4\pi R_S^2 \\
&= 4\pi \left(R_S^2 + \frac{4Gh}{c^3} + \frac{4G^2h^2}{c^6 R_S^2}\right) - 4\pi R_S^2 \\
&= \frac{16\pi Gh}{c^3} \\
&= 16\pi \ell_p^2
\end{aligned} \tag{4.18}$$

where the term of order $\frac{G^2h^2}{c^6 R_S^2}$ is negligible and ℓ_p^2 is the square of the Planck length [71]. Hence up to a factor of 16π the area of a black hole increases by one Planck area every time one bit of information falls into it. This implies that the area of a black hole horizon measured in Planck units is proportional to its entropy (hidden information) measured in bits: this very basic calculation therefore implies (up to a numerical factor) Bekenstein's famous formula $S = \frac{A}{4}$ [72]. As Susskind memorably puts it: "*information equals area*" [28]; this link between information, entropy, and area hints at the holographic principle [73–75] and AdS/CFT duality which we will look in section 5.2.9.

Chapter 5

The Information Paradox

5.1 An Exposition of the Information Paradox

(In this section we have set $c = \hbar = G = k_B = 1$.)

The black hole information paradox can be stated as follows:

Consider some matter in a pure state that collapses to form a black hole. After the hole has evaporated away completely, leaving only thermal Hawking radiation, the result of the whole process will have been that an initially pure quantum state evolved into a mixed state. This pure-to-mixed state process is not unitary and so violates a central principle of quantum mechanics.

Quantum mechanics requires that an initial state evolves according to an S-matrix: $|\Psi_{final}\rangle = S|\Psi_{initial}\rangle$. The unitarity of the S-matrix implies that the evolution is deterministic [3] and it indicates that the inital state can be retrieved from the final state according to $|\Psi_{initial}\rangle = S^{\dagger}|\Psi_{final}\rangle$. For example if all of the Hawking radiation were to be collected this should allow us to calculate what initial matter formed the hole. N.B. the relationship between information loss and pure-to-mixed state quantum evolution is subtle, and one does not necessarily imply the other, we will define these terms and their interplay properly at the end of this section.

Before we delve into the information paradox let's look at the closely related conundrum of what *Mathur* [5] calls "the entropy puzzle": what are the microstates of the black hole that account for its enormous entropy? We saw at the end of section 4.2 that the

Bekenstein entropy of a black hole is equal to a quarter of its area expressed in Planck units

$$S = \frac{A}{4} \tag{5.1}$$

and that by combining this entropy with the amount external to the hole the second law of thermodynamics still holds. But where does the Bekenstein entropy come from?

From statistical mechanics we think of entropy as emerging from a logarithmic function of the number of different states making up the system per macrostate. This would lead us to believe that the number of microstates per macrostate in a system with entropy S would be of order e^S. For a relatively small black hole, of solar mass, the number of microstates would be of order $10^{10^{77}}$. From what we know about black holes it seems like they should have very few microstates – they have "no hair" – a general black hole can be fully described by its mass, charge, and angular momentum [12, 76]. What could these microstates be? Here are some suggestions by various authors [77]:

- Black hole microstates may be the various combinations of internal matter and gravitational states making up a hole of given mass, angular momentum, and charge [78].
- There is a gas of quanta just outside of the event horizon whose entropy is that of the Bekenstein entropy of the hole [79, 80].
- The entropy may derive from the entanglement between quantum fields inside and outside of the horizon. The whole state taken together may be pure and so the entanglement entropy is found by tracing out the internal states [81, 82].
- Black hole entropy is the Noether charge associated with diffeomorphism symmety in theories containing higher curvature terms than classical general relativity. This can lead to corrections to the entropy formula $S = \frac{A}{4}$ [83].
- Microstates are different string excitations in string theory [84]. We will look into this in section 5.2.10.

There have also been attempts within loop quantum gravity which have correctly red-erived the black hole entropy (once a parameter has been fixed) [85, 86].

Let's now look at particle production by black holes in detail to see exactly why the process seems to evolve an initially pure state into a final mixed state. The black holes we'll treat in this section will be the Schwarzschild type we introduced earlier on. This section follows closely the treatment given in Mathur [87] as it clearly illustrates the robustness of the problem and how small corrections to the Hawking radiation do not present a solution.

In our present analysis of black hole evaporation we'll take certain "nice" spacelike slices (which we will soon define properly) through the spacetime containing the black hole (see Figure 5.1) which penetrate the event horizon and allow us to analyse the properties of the Hawking pairs – specifically their entanglement – which is very important when assessing their information encoding characteristics and their susceptibilities to small perturbations. In the following we will assume that we are working in the semiclassical domain, where quantum gravity effects are unimportant, and so our spacelike slices obey certain "niceness" conditions [87]:

- The quantum states we are studying should be contained entirely on a spacelike slice whose intrinsic curvature $R^{(3)}$ is everywhere smaller than the Planck scale: $R^{(3)} \ll \frac{1}{\ell_p^2}$.
- The spacelike slice should be embedded in 3+1 dimensional spacetime with small extrinsic curvature: $K \ll \frac{1}{\ell_p^2}$.
- In the neighbourhood of the spacelike slice, the four-curvature of the spacetime should also be small: $R^{(4)} \ll \frac{1}{\ell_p^2}$.
- Any matter present on the slice should not approach the Planck scale where quantum gravity effects are expected to apply.
- States on one slice will be evolved smoothly to another nice slice at a later time.

When our nice spacelike slices we have defined above deform, this produces Hawking pairs on the slices. The pairs have wavelengths of the order of the curvature length scale of the deformation, for a generic black hole this scale would be the Schwarzschild radius. This is analogous to our earlier discussion of particle pair creation in a changing harmonic potential at the end of section 3.1. Let the state of the produced entangled pair be of the simple form [87]:

$$|\psi\rangle_{pair} = \frac{1}{\sqrt{2}}|0\rangle_c|0\rangle_b + \frac{1}{\sqrt{2}}|1\rangle_c|1\rangle_b \tag{5.2}$$

where the b particles fly out to infinity (detected as Hawking radiation) and the c states fall in towards the singularity. The matter $|\psi\rangle_{matter}$ which forms the black hole is also contained on the spacelike slice but can be thought of as being far enough away that it has negligible effect on the Hawking pairs.

The whole quantum state $|\Psi\rangle$ describing the matter making up the hole and the pair is therefore given by:

$$|\Psi\rangle \approx |\psi\rangle_{matter} \otimes \left(\frac{1}{\sqrt{2}}|0\rangle_c|0\rangle_b + \frac{1}{\sqrt{2}}|1\rangle_c|1\rangle_b \right) \tag{5.3}$$

where the equality is only approximate as the matter may have a very small effect on the pair. Setting $|\psi\rangle_{matter}$ to be a two-level quantum system, for example

$$|\psi\rangle_{matter} = \frac{1}{\sqrt{2}}|\uparrow\rangle_{matter} + \frac{1}{\sqrt{2}}|\downarrow\rangle_{matter} \tag{5.4}$$

would give

$$|\Psi\rangle \approx \left(\frac{1}{\sqrt{2}}|\uparrow\rangle_{matter} + \frac{1}{\sqrt{2}}|\downarrow\rangle_{matter} \right) \otimes \left(\frac{1}{\sqrt{2}}|0\rangle_c|0\rangle_b + \frac{1}{\sqrt{2}}|1\rangle_c|1\rangle_b \right). \tag{5.5}$$

Quantifying the small effect $\epsilon \ll 1$ of the matter on the pair allows us to promote our above formula from an approximate identity into a full identity and gives [87]:

$$|\Psi\rangle = \left(\frac{1}{\sqrt{2}}|\uparrow\rangle_{matter} + \frac{1}{\sqrt{2}}|\downarrow\rangle_{matter} \right) \otimes \left(\left(\frac{1}{\sqrt{2}} + \epsilon \right) |0\rangle_c|0\rangle_b + \left(\frac{1}{\sqrt{2}} - \epsilon \right) |1\rangle_c|1\rangle_b \right). \tag{5.6}$$

The above perturbation is allowed, however the following change of state is not allowed [87]:

$$|\Psi\rangle = \left(\frac{1}{\sqrt{2}}|\uparrow\rangle_{matter}|0\rangle_c + \frac{1}{\sqrt{2}}|\downarrow\rangle_{matter}|1\rangle_c \right) \otimes \left(\frac{1}{\sqrt{2}}|0\rangle_b + \frac{1}{\sqrt{2}}|1\rangle_b \right) \tag{5.7}$$

as it drastically changes the entanglement properties of the system. The entanglement entropy of a subsystem was given in equation (4.15); what are the entanglement entropies for the above systems?

Tracing over $|\psi\rangle_{matter}$ and c for (5.5) gives $S_b = \ln 2$, for (5.6) it gives $S_b = \ln 2 - \epsilon^2 (6 - 2\ln 2)$, and finally for (5.7) $S_b = 0$ [87]. The entanglement entropy for the Hawking pair slightly affected by the matter $|\psi\rangle_{matter}$ can be seen above to tend quickly to the unperturbed value $\ln 2$ as ϵ is brought to zero. The final zero-valued entanglement entropy cannot be caused by the influence of faraway $|\psi\rangle_{matter}$ because locality on the

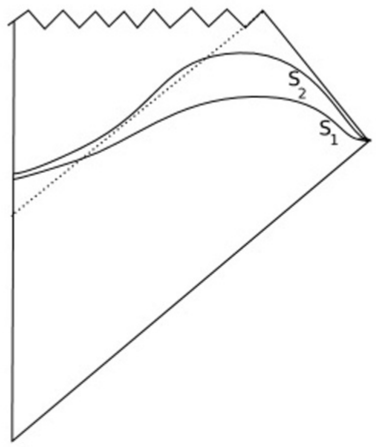

FIGURE 5.1: Nice slices through a black hole spacetime. Illustration from [87].

spacelike slice doesn't allow it. So locality implies:

$$\left| \frac{S_b}{\ln 2} - 1 \right| \ll 1. \tag{5.8}$$

An essential facet of the evaporating black holes under study is that the regions around their horizon must not contain information about the holes (equivalent to the observation that an infalling observer experiences nothing strange whilst passing through the horizon – ignoring tidal effects). Specifically any point on the horizon has a neighbourhood on which quantum fields (neither trans-Planckian nor larger in wavelength than the horizon radius) evolve according to the semiclassical evolution of fields in empty curved spacetime [87]. We saw earlier that *empty* spacetime is relative to the observer, however in the case of evaporating black holes the curvature scale is of order Schwarzschild radius and so different observers may detect a different quantity of Schwarzschild radius wavelength quanta, but only to order 1 – for quanta of wavelength much less than Schwarzschild radius (but not below the Planck length) all observers will agree that the state is a vacuum, with little error [77].

For evaporating black holes the Schwarzschild solution (2.1) is valid – despite its time independence – as the evaporation takes a very long time (much longer than the current age of the universe). Therefore the solution applies at any chosen point during evaporation until the black hole's radius diminishes to the Planck length [87]. Taking nice spacelike slices through a spacetime containing a black hole must avoid the hole's central singularity so as to satisfy the niceness conditions defined above. Let's rigorously define these slices by splitting them up into parts outside, inside, along with a connecting piece across the horizon [77, 87]:

Inside horizon Σ_I: $\frac{M}{2} < r < \frac{3M}{2}$ and $r =$ constant. Can be smoothly connected to Schwarzschild origin $r = 0$ at times before singularity formation.

Outside horizon Σ_O: $r \gg 4M$ and $t =$ constant.

Connection Σ_C: Smoothly connects Σ_O and Σ_I across the event horizon. Both space and time dimensions of Σ_C are of order M.

The spacelike slice defined above may seem strange as outside of the hole it is parameterized by $t =$ constant (as expected) whereas inside the hole it is parameterized by $r =$ constant. This is purely as a result of the Schwarzschild metric's pathology at the horizon where time and space coordinates swap roles.

A whole spacelike slice $\Sigma(t, r, C)$ is given by the union $\Sigma_I \cup \Sigma_O \cup \Sigma_C$. Shifting forward the parameters smoothly evolves one slice into another, i.e.

$$\Sigma_1 = \Sigma(t_1, r_1, C_1) \to \Sigma_2 = \Sigma(t_2, r_2, C_2) = \Sigma(t_1 + \delta t, r_1 + \delta r, C_1 + \delta C). \tag{5.9}$$

When evolving forward the spacelike slices the geometry of Σ_C can be taken to be unchanged as long as $\delta r \ll M$ [77]. This results in the Σ_I segments getting longer as Σ_O shifts forward in time – see Figure 5.2 for a simple diagram of this stretching. A succession of these nice spacelike slices builds up an entire spacetime containing the black hole. Evolving the slices along a timelike normal will leave the instrinsic geometries of both Σ_O and Σ_I (although it stretches) the same; however the connecting part Σ_C must grow in order to connect these segments together and to account for the longer Σ_I at successive intervals – this stretching happens only in the neighbourhood of Σ_C with spatio-temporal dimensions of order M [77, 87].

This stretching of Σ_C is the cause of the Hawking pair generation and we can now see why it can only occur in the presence of a black hole and not in Minkowskian spacetime. In flat spacetime Σ_C would eventually become null and then timelike, but due to the swapping of time and space coordinates at event horizons our choice of slices in black

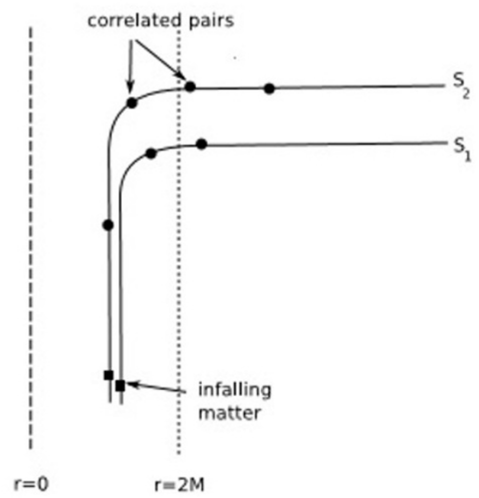

FIGURE 5.2: The stretching of nice slices during black hole evolution. Illustration from [87].

hole backgrounds always stay spacelike throughout their evolution and Hawking pair production carries on unimpeded until the hole reaches Planckian dimensions [77, 87].

Now let's look at how the entanglement of Hawking pairs changes as the black hole continually evaporates. This is key in understanding how Hawking radiation differs from thermal radiation emitted normally by objects such as a burning encyclopedia or a piece of coal; these objects emit radiation in a *fundamentally* different way that allows information conservation and retrieval in principle. A series of nice spacelike slices as time increases:

Slice 1: Collection of matter $|\psi\rangle_{matter}$ exists on slice but hasn't yet collapsed to form black hole.

Slice 2: Black hole has now formed. "Middle" of slice stretches to create first Hawking pair à la (5.3):

$$|\Psi\rangle \approx |\psi\rangle_{matter} \otimes \left(\frac{1}{\sqrt{2}} |0\rangle_{c_1} |0\rangle_{b_1} + \frac{1}{\sqrt{2}} |1\rangle_{c_1} |1\rangle_{b_1} \right), \qquad (5.10)$$

giving entanglement entropy ln2 between the Hawking radiation and the rest of the system as we saw previously when we assumed $|\psi\rangle_{matter}$ was a simple two-level quantum system (5.4).

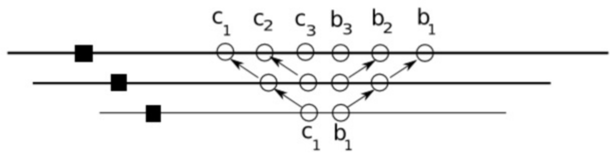

FIGURE 5.3: Hawking pair creation showing "pushing" of earlier ones away from the horizon as new pairs emerge. $|\psi\rangle_{matter}$ shown as black boxes on left of figure. Illustration from [87].

Slice 3: The matter $|\psi\rangle_{matter}$ is unchanged, the b_1, c_1 pair gets "pushed" outwards – see Figure 5.3 – by the stretching of Σ_C which also stimulates creation of a new Hawking pair b_2, c_2 giving full quantum state on the slice:

$$\begin{aligned}|\Psi\rangle \approx |\psi\rangle_{matter} &\otimes \left(\frac{1}{\sqrt{2}}|0\rangle_{c_1}|0\rangle_{b_1} + \frac{1}{\sqrt{2}}|1\rangle_{c_1}|1\rangle_{b_1}\right) \\ &\otimes \left(\frac{1}{\sqrt{2}}|0\rangle_{c_2}|0\rangle_{b_2} + \frac{1}{\sqrt{2}}|1\rangle_{c_2}|1\rangle_{b_2}\right)\end{aligned} \tag{5.11}$$

giving entanglement entropy 2ln2 between the Hawking radiation and the rest of the system.

Slice N+1: At the limit of the (N+1)-th slice N Hawking pairs have been created:

$$\begin{aligned}|\Psi\rangle \approx |\psi\rangle_{matter} &\otimes \left(\frac{1}{\sqrt{2}}|0\rangle_{c_1}|0\rangle_{b_1} + \frac{1}{\sqrt{2}}|1\rangle_{c_1}|1\rangle_{b_1}\right) \\ &\otimes \left(\frac{1}{\sqrt{2}}|0\rangle_{c_2}|0\rangle_{b_2} + \frac{1}{\sqrt{2}}|1\rangle_{c_2}|1\rangle_{b_2}\right) \\ &\qquad\qquad\qquad\qquad\qquad\qquad ... \\ &\left(\frac{1}{\sqrt{2}}|0\rangle_{c_\mathrm{N}}|0\rangle_{b_\mathrm{N}} + \frac{1}{\sqrt{2}}|1\rangle_{c_\mathrm{N}}|1\rangle_{b_\mathrm{N}}\right)\end{aligned} \tag{5.12}$$

giving entanglement entropy Nln2 between the Hawking radiation and the rest of the system.

Now the problem is clear. Once the black hole has emitted all of its mass in Hawking radiation it ceases to exists, leaving only a radiation field: represented by $b_1, b_2, ..., b_\mathrm{N}$ above. This final Hawking radiation has entanglement entropy Nln2 but it is not *entangled with anything*. Therefore the radiation can only be described by a density matrix, it is now a mixed state [77]. Before the black hole had evaporated away completely the Hawking radiation was entangled with internal states of the hole and so the whole quantum system had zero von Neumann entropy and was pure.

What is the effect of the Hawking pairs on each other? Could these corrections be sufficient to encode the information about the matter making up the black hole in correlations in the outgoing radiation? It can be shown [77, 87] that as each new Hawking pair is created the entanglement entropy of the outgoing radiation increases steadily, taking into account possible corrections, by an amount ln2-2ϵ per pair ($\epsilon \ll 1$ as before). This increase in entropy is the minimum possible increase whenever *any* Hawking pair is created. Therefore Hawking's original derivation [24] is sturdy in the face of small corrections and something more drastic is needed to solve the information paradox.

Although evolution of a pure to a mixed state violates unitarity, this does not always imply information loss – as emphasised in Mathur [87] in the context of black hole evaporation. In some cases a mixed Hawking radiation field can contain all of the information about the matter that had formed the hole, and on the other hand a final pure state could result in information loss. Let's look at some examples, firstly a process which violates unitarity but conserves information. Starting with a quantum matter state:

$$|\psi\rangle_{matter} = \alpha|1\rangle_{matter} + \beta|0\rangle_{matter}. \tag{5.13}$$

After this matter has collapsed to form a black hole and two Hawking pairs (b_1, c_1) and (b_2, c_2) have emerged from the vacuum, the complete quantum state of the system becomes:

$$\begin{aligned} &\left(\frac{1}{\sqrt{2}}|1\rangle_{matter}|0\rangle_{c_1} + \frac{1}{\sqrt{2}}|0\rangle_{matter}|1\rangle_{c_1}\right) \\ &\qquad \otimes (\alpha|1\rangle_{b_1} + \beta|0\rangle_{b_1}) \\ &\otimes \left(\frac{1}{\sqrt{2}}|1\rangle_{c_2}|0\rangle_{b_2} + \frac{1}{\sqrt{2}}|0\rangle_{c_2}|1\rangle_{b_2}\right). \end{aligned} \tag{5.14}$$

In the above case the first Hawking radiation particle b_1 contains *all* of the information regarding the matter that collapsed to form the hole. This can be retrieved by an observer and information in principle is now not lost. However the second Hawking radiation particle b_2 is entangled (with entanglement entropy ln2) with its partner c_2 within the hole; therefore after the hole has completely evaporated the final state is mixed and we have violated unitarity of quantum evolution.

What about a system with the opposite problem? Letting our original matter state (5.13) now evolve into a black hole that has produced so far one Hawking pair (b, c) given by:

$$\begin{aligned}(\alpha|1\rangle_{matter}|0\rangle_c + \beta|0\rangle_{matter}|1\rangle_c)\\ \otimes\left(\frac{1}{\sqrt{2}}|1\rangle_b + \frac{1}{\sqrt{2}}|0\rangle_b\right)\end{aligned} \tag{5.15}$$

results in a Hawking radiation particle b existing in a pure state but carrying no information about the original matter that we are interested in.

N.B. the two evolutions we just looked at are toy models and cannot result from normal semiclassical processes in real black hole evaporation [87]. Normal black hole evaporation suffers from *both* of issues shown above – non-unitarity and information loss – and a satisfactory solution to the information paradox would need to find a remedy for both.

Why is information loss bad? One reason is that in theories containing the possibility of information loss, energy seems to be unconserved; Preskill suggests a heuristic picture of this process as analogous to the coupling of a signal to a source of random noise [88]. A quantum state in the universe can be thought of as a signal encoding information, if the universe is then coupled to a source of noise it can overwhelm the signal destroying the information and at the same time pump energy into the universe, violating energy conservation. Another reason why information loss is undesirable is that, if violated, quantum theory would have to be drastically altered if not completely overhauled, and due to quantum theory's great successes and continued experimental validations this would seem unreasonable.

5.2 Proposed Solutions

In this section we shall look at a number of potential solutions to the information paradox that have been put forward in the last few decades. Firstly taking a cursory look at some of the less promising and less discussed candidate solutions before going into more detail with regards to complementarity, the AMPS paradox, and the fuzzball conjecture.

Several other proposed solutions not covered here can be found in *Mathur* [89].

5.2.1 Remnants

The point where the black hole vanishes seems to cause a violent shift from pure to mixed state. Is there a way to terminate the evaporation before the black hole completely disappears?

The Hawking evaporation process could cease when the hole reaches the Planck mass, leaving what is known as a remnant [90]. This would leave unitarity intact but such remnants would behave unlike any other known objects; as the remnant would be entangled with entropy Nln2 with the emitted radiation it must have at least N possible internal states [87] but be of finite size and energy. As N is arbitrary and not bounded the remnant can have an abitrarily large degeneracy unlike normal quantum states, this feature can also lead to divergences when formulating its interaction with normal matter [77].

5.2.2 Bleaching

Bleaching posits that the information in a quantum system is somehow prevented from entering the black hole – the horizon "bleaches" it – possibly decoupling the information from the system's energy and momentum allowing it to come out in Hawking radiation [30, 77]. This would require a drastic alteration of the semiclassical assumptions of the emptiness of the horizon that we assumed earlier when looking at Hawking pair entanglement. The equivalence principle would be violated as a physical process must occur at the horizon to decouple the information or to prevent the system from falling in.

5.2.3 Quantum Hair

Could the information about infalling matter be contained in distortions of the event horizon? This contradicts the "no-hair theorems" [76] and attempts to create "hair" from quantum fields around the horizon have been shown to lead to divergent energy tensors at the horizon [91]. Some of the attempts to derive quantum hair have utilised discrete gauge symmetries, while this somewhat circumvents the no-hair theorems it isn't sufficient to explain the enormous degeneracy implied by the Bekenstein entropy formula [92, 93]. The fuzzball conjecture predicts quantum hair can be formed from bound states of strings and branes which we will look at in more detail in section 5.2.10.

5.2.4 Baby Universes

This idea proposes that the matter normally thought of as collapsing to a singularity during black hole formation actually generates a "baby universe" in which the matter then resides, caused by quantum gravitational effects [30]. This new universe is causally disconnected from our own and so the information is lost from our vantage point. However this new universe can be thought of as one section of a *multiverse* (another section obviously being our own universe). From the point of view of an observer capable of making measurements on the whole multiverse information is not destroyed, just transported. Therefore from our point of view the black hole seems to have evolved from a pure to a mixed state, whereas in actuality our universe is just a subsystem which becomes pure when combined with the baby universe subsystem [30, 94].

A deeper analysis of this idea actually shows that black hole evaporation does lead to a final pure state in our universe [95], but we are still left with the problem of the unknown mechanism by which the Hawking radiation encodes the information regarding the collapsed matter, therefore the baby universe hypothesis seems to lack important elements needed to solve the paradox satisfactorily.

5.2.5 Information Emergence at the End of Evaporation

What about if the information leaks out of a black hole right at the end of its evaporation process, when the hole is of Planck scale?

Assuming a black hole follows the normal semiclassical evaporation process until its radius is around the Planck distance, then this takes time roughly of the initial black hole mass cubed M^3 [30]. But then the time required for the Planck-dimensional hole to completely disappear is of order $t \sim M^4$ as we now show. Assuming the remnant emits quanta, the number of which is of order Bekenstein entropy $S \sim M^2$, and taking the energy of the remnant to be of order one $E \sim 1$, then each quantum emitted has energy of order M^{-2} with wavelength given by inverse of the energy. For the quanta to encode the information that was contained in the remnant then the quanta must emerge with minimal overlap; given that the time taken per quantum emission is $\sim M^2$ and there are $\sim M^2$ quanta to come out we arrive at a *minimum* time required for the Planck-dimensional hole to disappear given by $t \sim M^4$ [30, 96]. This leaves us with very long-lived remnants of the type we looked at in 5.2.1, therefore we are burdened with the same problems we encountered there, namely the potentially infinite degeneracy of the remnants.

Recent work [97] has shown that even if we postpone information emergence until after the black hole has reached Planckian dimensions this does not imply that the remnant contains all of the information. This counterintuitive result is due to the fact that information isn't additive, a small amount of quanta can "lock" (render inaccessible) a large amount of information, with the information only emerging when the hole has *completely* evaporated [90].

5.2.6 Treating Black Holes as Rubik's Cubes

Recently a proposal was put forward to treat black holes as a collection of "Rubik's cubes", configurations of which represent the microstates of a black hole. Let's look at this proposal in more detail before explaining its flaws [89, 98].

For simplicity a Rubik's cube is represented by a 2-by-2 grid of numerals. A possible grid being:

2	1
3	4

.

These numerals span the internal Hilbert space of the black hole (the black holes we are looking at have zero angular momentum and charge), their combinations in grids representing orthogonal states, with the vacuum state represented by

1	2
3	4

.

Czech et al [98] also define four operators: $\overleftarrow{L}$ which swaps the numerals in the first column, $\overleftarrow{R}$ swaps numerals in the second column, $\overleftarrow{U}$ swaps numerals in the top row, and $\overleftarrow{N}$ leaving states unchanged. The Hawking pairs created by a black hole are split into ingoing and outgoing particles as we have seen. In this model the ingoing Hawking particles act as operators on the internal state of the black hole with the accompanying outgoing particles – contingent on this operation – being of four possible types n, l, r, u. For clarification: if during the Hawking process an ingoing particle represented by operator $\overleftarrow{R}$ is created then the internal state of the hole will change (the numerals in the right column will swap) and an outgoing Hawking radiation particle will be emitted in state r, the other operators working in analogous ways. The $\overleftarrow{N}$ operator will leave the internal state unchanged and emit an "n-particle" which is in actuality the absence of a particle. An appropriate unitary operator acting on the internal state of the hole and the state describing the previously emitted Hawking radiation is given by [98]

$$\overleftarrow{S} = \frac{1}{2}\left(\overleftarrow{N} \otimes n + \overleftarrow{L} \otimes l + \overleftarrow{R} \otimes r + \overleftarrow{U} \otimes u\right). \tag{5.16}$$

In order to more realistically model the black hole, a *collection* of E number of grids is used to represent internal state of the hole (the letter E is suggested by *Czech et al* [98] as the variable is designed to be redolent of the entropy, energy, or inverse temperature of the hole as it evaporates, i.e. E decreases until the hole is no more). When the evaporation operator S "solves" a grid (evolves it to the vacuum configuration), the grid is removed from the state and a particle is emitted representing a drop in energy of the hole (detectable by an outside observer) [98]. After the unitary operator S has acted on the hole many times (analogous to the hole evaporating over a long time period) then the number of "unsolved" grids E decreases and the entanglement entropy between outside and inside states of the hole drops to zero at the point of evaporation. Therefore this model seems to describe an evaporating black hole losing all of its energy to pure Hawking radiation, seemingly without information loss or pure-to-mixed state evolution. Is this model physical?

Unfortunately the Rubik's model is not a physical analogue of Hawking radiation. The evolution of the system given in *Czech et al* [98] cannot be physically realised by a hole well-described by the Schwarzschild metric [89] – the no-hair theorems show that this metric cannot be significantly deformed [76] – and is in fact a model of a generic burning object. This model could be used to describe a burning encyclopedia, in which case the entanglement entropy does drop to zero after it has been completely burnt leaving a pure radiation field, however black hole evaporation creates a steadily increasing entanglement entropy as we saw earlier which this model does not address [89].

5.2.7 Tunnelling

Research has been done [77, 89, 99, 100] treating the black hole evaporation process as a case of *quantum tunnelling*. Tunnelling is normally thought of in the context of a classically forbidden transition of a particle from one side of a potential barrier to another; in the black hole context the Hawking radiation particle is thought of as travelling from the inside to the outside of the hole, another classically forbidden process. *Parikh and Wilczek* [99] showed that Hawking radiation can be derived using tunnelling considerations but that the emission spectrum deviates from a thermal spectrum, the authors believed this to be suggestive of the radiation encoding information within its non-thermal correlations. However as pointed out in *Mathur* [89] the non-thermal spectrum has no direct bearing on the information paradox as the entanglement between outside and inside Hawking modes still increases steadily as we saw in section 5.1.

Studying Hawking radiation in the context of tunnelling does have certain advantages however. It is one of a handful of methods that can be used to investigate fermionic spin-1/2 emission (and that of higher spins) [77, 101] from black holes; and provides an independent check of their thermodynamical properties, including the Hawking temperature and more generally the Unruh temperature in Rindler spacetimes [101].

5.2.8 Complementarity, Firewalls, and ER=EPR

Black hole complementarity is a very intriguing idea, introduced in the early 1990s, based on the proposition that a physical theory doesn't need to be able to describe an observer who can measure phenomena both inside and outside of a black hole because such an observer cannot exist [36, 102].

Complementarity attempts to get around the information paradox by positing that matter falling into a black hole, rather than passing through unimpeded, hits a "stretched horizon" located a Planck distance above the Schwarzschild horizon, which absorbs and unitarily reemits it as Hawking radiation [103]. As we saw in section 2.1 the Schwarzschild coordinates (2.1) seem to show that an observer falling towards a black hole would take an infinite amount of time to reach the horizon, this was taken as evidence that the Schwarzschild time coordinate t was unphysical. Complementarity implies that infalling matter indeed never falls into the hole *and* passes through *at the* *same time* without implying a contradiction. In other words after throwing a bit of information into a black hole, an infalling observer sees it inside whereas an external observer sees the same bit outside, *but there is no one who can see both* [104]. The word *complementarity* is used because just as in the Bohrian concept of complementarity (where the wave and particle properties of light cannot be measured simultaneously) trying to measure a bit of information inside a black hole precludes being able to measure it outside and vice versa, therefore there is no contradiction as no one observer can achieve both [28]. Black hole complementarity assumes that information emerges in Hawking radiation, so what would happen if a bit of information fell into a black hole and an external observer waited long enough to decode the radiation and retrieve the bit? Could they then carry that bit into the hole and compare it with the bit (the same bit) that fell in? This would effectively break down the complementarity framework and also imply that the black hole acted as a *cloning machine* (forbidden by quantum information theory); in actual fact this scenario is prohibited as the time it would take for an external observer to gather and decode the bit from the Hawking radiation would be much longer than the time taken for the infalling bit to reach and be destroyed by the singularity [28, 63, 105].

The postulates underlying black hole complementarity are [36, 104]:

- A distant, external observer describes the formation and evaporation of a black hole as unitary, i.e. in principle there exists a unitary S-matrix describing the whole process.
- Outside of the stretched horizon of the black hole (at distances $> R_S + \ell_p$) physics can be described using semiclassical field theory.
- The subspace of states describing a hole has dimensions of e^S, where S is the Bekenstein entropy: $S = \frac{A}{4}$.
- To an infalling observer the horizon seems like any other region of spacetime – in other words the equivalence principle holds.

Let's look at the original reasoning of *Susskind et al* [36] which led to the idea of black complementarity.

The Penrose diagram for an evaporating black hole is given in Figure 5.4. Assuming the spacetime in which the black hole lives is globally hyperbolic we can foliate it using a series of Cauchy surfaces – cf. section 3.1. A Cauchy surface for an evaporating black hole can be defined which is partially inside and outside of the hole, split into Σ_{bh} and Σ_{out} in Figure 5.4, where $\Sigma_{\mathrm{P}} = \Sigma_{\mathrm{bh}} \cup \Sigma_{\mathrm{out}}$ – the Cauchy surfaces we are looking at satisfy the niceness conditions defined in section 5.1. (The Hilbert space of states on Σ_{P} is a product state of Hilbert spaces defined for the inside and outside of the black hole [36].) The last of the niceness conditions says that states defined on one slice can be evolved smoothly to states on slices at later times, therefore a state $|\psi(\Sigma)\rangle$ on Σ can be evolved smoothly to a state $|\psi(\Sigma_{\mathrm{P}})\rangle$ on Σ_{P} which can then further evolve to $|\psi(\Sigma')\rangle$ defined on Σ' (this last space-like slice describes the universe after the black hole has fully evaporated). As stated above, the first postulate of black hole complementarity says that $|\psi(\Sigma')\rangle$ evolves from $|\psi(\Sigma)\rangle$ by a unitary S-matrix – also $|\psi(\Sigma')\rangle$ must be a pure state as we assume that the Hawking radiation is pure after complete black hole evaporation [36]. It is also assumed that the state $|\psi(\Sigma')\rangle$ evolves smoothly from a state $|\xi(\Sigma_{\mathrm{out}})\rangle$ defined on Σ_{out} [36]. This implies (along with the fact that the Hilbert space on Σ_{P} is a product space) that $|\psi(\Sigma_{\mathrm{P}})\rangle = |\Pi(\Sigma_{\mathrm{bh}})\rangle \otimes |\xi(\Sigma_{\mathrm{out}})\rangle$ where $|\Pi(\Sigma_{\mathrm{bh}})$ is a state defined on Σ_{bh}.

The product state defined on slice Σ_{P}, $|\psi(\Sigma_{\mathrm{P}})\rangle$, evolves linearly from $|\psi(\Sigma)\rangle$ however $|\xi(\Sigma_{\mathrm{out}})\rangle$ also evolves linearly from $|\psi(\Sigma)\rangle$, therefore it seems to follow that $|\Pi(\Sigma_{\mathrm{bh}})\rangle$ *does not depend on* the pre-black hole formation initial state $|\psi(\Sigma)\rangle$ – the event horizon seems to *bleach* (see section 5.2.2) the information of an infalling state [36, 77]. Complementarity removes this bleaching effect by arguing that the assumption that there exists a state Σ_{P} simultaneously describing the inside and outside of a hole is unphysical and so the above argument collapses [36].

N.B. the stretched horizon of the black hole must be “virtual” as an infalling observer does not measure it and complementarity doesn’t provide any hair with which a stretched horizon could thermalise and unitarily reemit the infalling information [77, 106].

As we have seen previously, if semiclassical physics is valid on nice slices through a black hole spacetime then Hawking pairs will inevitably emerge from the vacuum *not* from a stretched horizon, therefore complementarity demands that some new non-local physics must be important over Schwarzschild horizon scales [77, 106]. A recent paper

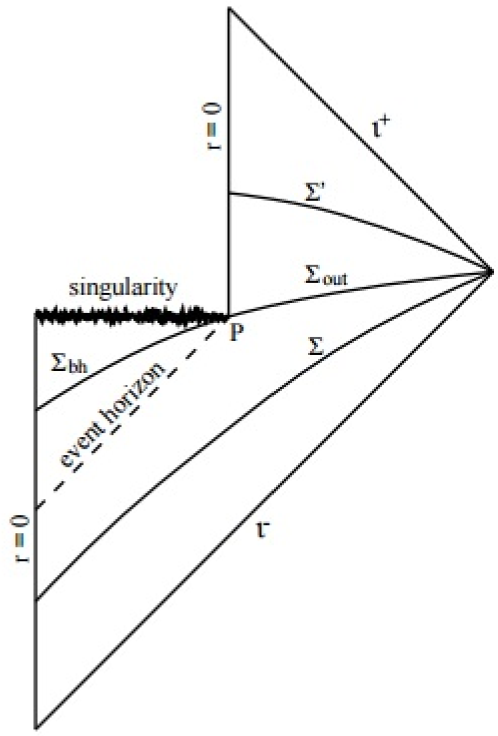

FIGURE 5.4: Penrose diagram for evaporating black hole foliated by space-like Cauchy surfaces. Past and future null infinites, singularity and event horizon also shown. Illustration from [36].

[37] – known as AMPS after its authors – argues that the postulates of complementarity lead to a contradiction, and that assuming information conservation and the validity of effective field theory away from the horizon implies the existence of a "firewall" at the horizon – a region of Planck density radiation that would destroy any infalling matter – thus doing away with the equivalence principle which many authors believe to be too radical a step [77, 104, 107].

The AMPS argument can be put as follows [106]. Assuming the tenets of complementarity to be correct, a faraway external observer (A) receives Hawking radiation emitted from a stretched horizon whereas an infalling observer (B) experiences nothing unusual as they fall into the hole. For observer A early Hawking radiation is located far away from the hole whereas recently emitted radiation is just outside the stretched horizon; observers A and B find no discrepancies between measurements of early and recent Hawking quanta [106]. Observer B measures a vacuum at the horizon and so as we saw at the end of section 5.1 the entanglement entropy of the hole with Hawking radiation must steadily grow as each new Hawking pair is created. However based on the assumption that observers A and B measure the same early- and late-time Hawking quanta, the entanglement entropy between the hole and the outgoing radiation also appears to steadily increase for observer A; this leads to a contradiction as observer A should measure the entanglement entropy to decrease after the hole is halfway through its evaporation in order for the final Hawking quanta field to be pure [106, 108]. By removing the assumption that observer B experiences nothing at the horizon this allows the information to be retrieved by observer A however observer B then is burnt up at the horizon by the firewall [37].

Complementarity demanded non-locality limited to Schwarzschild horizon scales however it assumed local physics exterior to the stretched horizon, the AMPS argument suggests that this limitation leads to a seemingly unphysical firewall [77, 106], therefore the original idea of complementarity [36] does not appear to provide a consistent mechanism by which information is conserved in black hole evaporation.

Much work has been done attempting to bypass the existence of firewalls (see *Mann* [77] for a summary of recent attempts) including string theoretic studies of black holes [26] (see section 5.2.10), and a very intriguing idea known as the ER=EPR proposal [109].

ER=EPR – the name deriving from the conjectured general equivalence of entanglement (cf. the Einstein-Podolsky-Rosen paradox) and wormholes (also known as Einstein-Rosen bridges) – suggests that outgoing Hawking particles are connected to their partners via wormholes [110]. Let's see how this may help with removing firewalls. Another way of stating the AMPS paradox is that an outgoing Hawking quantum has to be entangled with all the previously emitted Hawking radiation if information is to be conserved, however the outgoing quantum is also entangled with its infalling partner as we have seen, both of these entanglements experienced by the outgoing quantum are *maximal* – meaning that making a measurement just on the outgoing quantum provides no information about the complete entangled state [111] – which is forbidden by the well-established *monogamy of entanglement* [107]. The AMPS authors did not want to sacrifice information conservation so they proposed that the entanglement between Hawking pairs had to be broken, thereby releasing a huge amount of energy at the horizon forming a firewall [37, 107]. The ER=EPR proposal says that the Hawking pairs are related by wormholes and are therefore not independent systems, allowing them to interact and leaving the entanglement between outgoing Hawking quanta untouched [109, 112]. The ER=EPR authors stress that the idea does not disprove the existence of firewalls however they state that "if it can be shown that the Einstein-Rosen bridge connecting the black hole to its radiation is smooth near the black hole, then there will be no firewall" [109] unless the outgoing radiation is manipulated in a very contrived way. This idea is very recent and is subject to much debate [112].

5.2.9 The AdS/CFT Correspondence

One of the most striking recent discoveries in physics is that string theory in a particular spacetime background is dual to a supersymmetric gauge theory on the boundary of that spacetime – the AdS/CFT correspondence [31, 33]. Before looking into the applicability of this duality to the information paradox let's introduce the AdS/CFT correspondence along with it most significant features.

The AdS/CFT correspondence emerged in the late 1990s [31] and is considered to be one of the most important and interesting results string theory has yet produced [113, 114]. In the simplest possible terms it posits a duality between gravitational theories and theories without gravity. The original formulation given by *Maldacena* [31] treated the equivalence of type IIB string theory compactified on $AdS_5 \times S^5$ and N = 4 supersymmetric Yang-Mills theory in 3+1 dimensions.

Definition: *Anti-de Sitter (AdS) space is a maximally symmetric solution of Einstein's field equations with a negative cosmological constant* [115].

Definition: *A conformal field theory (CFT) has no dimensional parameters and is generally also scale invariant (this last point is subtle, see last two of references at end of this definition). A CFT is particularly useful as it exists at a fixed point of the renormalization group's flow and one can easily study its geometry, i.e. define its metric.* [3, 116–118].

Definition: *$AdS_5 \times S^5$ denotes the ten-dimensional product space of five-dimensional AdS space and 5-sphere S^5* [3].

Definition: *$N = 4$ supersymmetric Yang-Mills theory is a quantum field theory with non-abelian gauge symmetry that relates bosonic and fermionic fields using four supersymmetries* [3, 119].

Definition: *Type IIB string theory describes closed, oriented superstrings of two types: left-moving and right-moving, which transform under separate supersymmetries having the same chirality* [53].

Since its discovery the correspondence has been applied to areas not originally thought to have been pertinent including condensed matter physics and relativistic hydrodynamics [116].

The AdS/CFT correspondence emerges from the duality between open and closed strings in string theory [113], however the duality can also be motivated by looking at lattice systems using renormalisation group methods following the approach given in *Ramallo* [116]. The renormalisation group treats the running of physical couplings at varying energy scales. The best known example of this phenomenon being the varying electron charge in QED [120]. The Hamiltonian of a system without gravity on a lattice, of spacing a, is given by [116]

$$H = \sum_{i,x} J_i(x, a)O^i(x) \tag{5.17}$$

where i numbers the different operators O^i, x parameterizes the lattice position, and $J_i(x, a)$ are the sources for the operators at each point on the lattice. By coarse-graining the lattice (increasing the spacing between lattice points and averaging the lattice variables) each operator O^i must be weighted differently, and the sources $J_i(x, a)$ must vary as the coarse-graining progresses [116]. For example, continually increasing the spacing between lattice points by a factor of four evolves the sources as

$$J_i(x, a) \to J_i(x, 4a) \to J_i(x, 16a) \to ... \tag{5.18}$$

and so on. As can be seen, during the above evolution the $J_i(x, a)$ are dependent on the scale $s = (a, 4a, 16a, ...)$. At weak coupling the variation of $J_i(x, a)$ with respect to the scale (i.e. the familiar β function) can be found using perturbation theory methods, however when the coupling becomes strong the AdS/CFT method suggests treating s as an additional dimension [116]. $J_i(x, a)$ are then defined as quantum fields in a space with one extra dimension with dynamics given by a particular metric (i.e. defining gravity), with the operators O^i defined on the *boundary* of this manifold. It may be thought of as strange that all of the bulk physics can be described by studying the space's boundary, however as we saw earlier the entropy of a black hole, all of its hidden information, is proportional to its area measured in Planck units, not its volume. Some authors have stated that the breakdown of local gravity theory at some scale due to the AdS/CFT duality may hold the key the solution of the information paradox [113].

What does a black hole look like in anti-de Sitter space? AdS space has constant negative scalar curvature creating an effective gravitational field that pulls objects to the center of the space, irrelevant of the mass distribution in the spacetime. A black hole formed in AdS space *won't evaporate*, the Hawking emission particles from the surface fall back to the surface [28].

If a black hole were created in AdS space this could, by Maldacena's duality [31], be equated with a CFT on the space's boundary. If the CFT were unitary then we would expect the bulk evolution to be unitary as well thereby conserving information and bypassing the information paradox; as stated in *Lowe and Thorlacius* [33]: the unitarity property of the boundary CFT "strongly suggests that all information about an initial state that forms a black hole is returned in the Hawking radiation". Does the AdS/CFT correspondence really provide a simple solution to the information problem or as *Lowe and Thorlacius* state does it just *strongly suggest* information conservation without providing the physical mechanisms for a satisfactory explanation?

As pointed out in *Mathur* [114], the statement that AdS/CFT proves information preservation in black hole evaporations is a circular argument. A black hole in AdS, just like in the Schwarzschild case looked at earlier, has a horizon where the state of quantum fields is the vacuum arising from normal gravitational collapse [89, 114], this vacuum state at the horizon is "stretched" during black hole evaporation creating entangled Hawking pairs. This evaporation process must result in either a breakdown of locality, pure-to-mixed state evolution, or a final remnant state [114]. As we saw at the end of section 5.1, the hole evaporation necessarily creates a steady increase of entanglement entropy between the interior and exterior of the hole, this result also holds in AdS space therefore the only way to ensure a pure final radiation state is to undermine some of the assumptions we made in our derivation of Hawking radiation in section 5.1 (we shall see which in particular in section 5.2.10). So we see that AdS/CFT does not solve the paradox. Black hole information conservation must be understood on the gravity side of the coin, not just by considering the CFT side.

5.2.10 Fuzzballs

Arguably the most successful potential solution to the information paradox is the "fuzzball" proposal [121–123].

In section 5.1 our derivation of Hawking radiation was explicitly predicated on the horizon having small spacetime curvature to leading order. Let's quantify the evolution of modes on the horizon, letting ψ_i be low-energy quanta (neither trans-Planckian nor larger in wavelength than the horizon radius) and evolving these forward in a time interval of order Schwarzchild radius of the hole gives matrix element:

$$\langle\psi_i|H|\psi_j\rangle. \tag{5.19}$$

Defining the matrix element for the semiclassical evolution of a quantum field on generic low-curvature spacetime as

$$\langle\psi_i|H_0|\psi_j\rangle \tag{5.20}$$

gives

$$\langle\psi_i|H|\psi_j\rangle = \langle\psi_i|H_0|\psi_j\rangle + O(\epsilon) \tag{5.21}$$

where ϵ is much less than unity. As pointed out in *Mathur* [114] if (5.21) is valid at the horizon of a black hole then evaporation will *inevitably* lead to information loss or remnants.

The fuzzball proposal undermines the horizon assumption (5.21), it states that there are corrections at *order unity* to the evolution of low-energy quanta at the horizon [114]. It also posits that black holes do indeed have "hair" (string configurations) which when taken into account can explain the origin of its large entropy, as suggested by *Susskind* [124]. Let's look now in more detail at how the fuzzball construction can explain the origin of black hole entropy and also potentially solve the information paradox. In the following we mainly follow the treatment of fuzzball construction laid out by *Mathur* [121] as well as several other sources cited below.

The information paradox relies on quantum gravity effects being confined to within a given distance – fuzzballs contradict this and undermine Hawking's argument. Hair wasn't found before as perturbative methods were used, the fuzzball constructions are nonperturbative and construct the hair required for unitary emission of Hawking radiation [121]. The fuzzball proposal emerges from string theory. Since string theory is thought to be complete and consistent then we must only use objects present within the theory when forming a black hole (i.e. strings, branes, 2-form fields, 4-form fields, and so on) [121].

A very important feature of string theory is *compactification*: whereby degrees of freedom are said to be compact or non-compact, allowing for dimensional reduction, generalising the beautiful work of Kaluza and Klein at the start of the last century [52, 125]. For a string existing in 9+1 dimensions, we can evidently compactify 6 of the spatial dimensions at such a scale that an observer (observer A) sees only 3+1 dimensions. Figure 5.5 shows an example of a dimension wrapped around a cylinder, invisible to observer A who cannot resolve this compactification. If we let a quantum of gravity, a graviton, travel around this compact dimension, observer A sees a point mass lying on the non-compact direction; this point mass then carries "momentum charge" n_p (if this charge is equal to the mass of the point seen by observer A then this is called a "BPS object") [121]. In a way analogous to above we can wrap a string around the compact dimension (around the cylinder), again this system appears as a point mass to observer A; now the point carries "winding charge" [5]. We now have some objects defined in string theory from which to construct a black hole.

The simplest construction is that of a "1-charge" black hole (using one type of object, i.e. a graviton *or* a wound string *or* some other objects we have yet to introduce) [26]. Firstly let's look at a 1-charge hole made from wound strings. Black holes normally studied are very massive and so to form a realistic hole we should not use a string wound around a compact direction once, but a string wrapped around many times (with its ends joined) in order to create an object with a higher winding charge n_1 and therefore higher mass (remember we are treating BPS objects: mass=charge) [121].

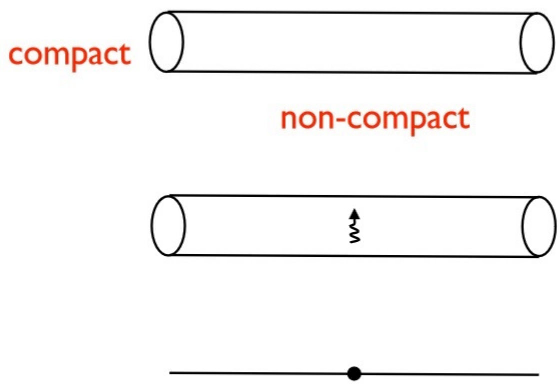

FIGURE 5.5: Top: Illustration of compactification. Middle: graviton travelling around compact dimension. Bottom: point mass appearance of graviton to observer A. Illustration from [121].

Counting the number of states of this string wound several times around the compact direction, taking into account supersymmetry, gives 256 configurations and so an entropy value given by the logarithm of this number [26, 121]. This entropy value is very small and more importantly it is *fixed*, it doesn't have a functional dependence on the winding charge n_1. The metric produced by this wound string gives a singular horizon and a zero-valued entropy, as expected from the count of states (ln256 is almost zero); another type of 1-charge hole, made from a high-energy graviton travelling around the compact direction, gives the exact same results as the wound string case [121].

What about a 2-charge hole? A wound string can maintain travelling waves (of momentum otherwise given by a graviton). This is called a NS1-P system, see Figure 5.6 – NS1 referring to the string and P to the momentum of the wave. The momentum of the travelling wave can be partitioned between different harmonics (i.e. all of the momentum could be put in the ground state or be more spread out amongst the modes) and these partitions give different configurations and thus their count can be used to find a microscopic value of the entropy of the system [121]. As can be seen pictorially in Figure 5.6 the total length of the NS1 string L_T is given by the circumference of the compact direction (thought of as a cylinder; c.f. Figure 5.5) multiplied by the winding charge n_1 [121]. A single k-th harmonic mode has momentum

$$p = \frac{2\pi k}{L_T}, \tag{5.22}$$

whereas the *total* momentum carried by the string is given by [91, 121]

$$p_{total} = \frac{2\pi n_p}{L} = \frac{2\pi n_1 n_p}{L_T}. \tag{5.23}$$

If each harmonic k has m_k modes then by equating two forms for the total momentum it trivially follows that

$$\sum_k k m_k = n_1 n_p. \tag{5.24}$$

The number of different configurations is of order [121]

$$e^{2\pi\sqrt{\frac{n_1 n_p}{6}}}, \tag{5.25}$$

partitioning these between c degrees of freedom (from the freedom to vibrate in the 8 spatial dimensions left available in the 9+1 dimensional system) alters the formula for the number of different configurations, now of order

$$e^{2\pi\sqrt{\frac{c n_1 n_p}{6}}}. \tag{5.26}$$

Exponentiating this number gives a value for the microscopic entropy (using supersymmetry to account for both bosonic and fermionic modes) [121] of

$$S_{micro} = 2\pi\sqrt{2 n_1 n_p}. \tag{5.27}$$

Let's put the above structures into the framework of type IIB string theory in order to take advantage of its dualities [53] and to find more complicated (3-charge, 4-charge, ...) constructions. First we compactify the spacetime thusly:

$$M_{9,1} \to M_{4,1} \times T^4 \times S^1, \tag{5.28}$$

where T^4 is a 4-torus and the NS1 and P of the 2-charge system we looked at previously will be contained on/around the circle S^1 [121]. Using both S- and T-dualities we find that for the NS1-P system, NS1 around S^1 is dual to a D5 brane wrapped in $T^4 \times S^1$, and the P charge becomes D1 around S^1 [52, 53, 121]. These dualities do not affect the non-compact directions which is useful as later we will see that the tranverse spreading of the modes is the source of the fuzzball's size; the dualities are only in the compact

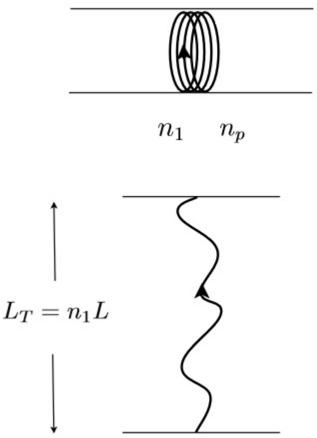

FIGURE 5.6: Top: 2-charge NS1-P system showing compact, and non-compact (left-to-right) directions. Bottom: Compact direction "rolled out" showing the string's full length. Illustration from [121].

directions and so do not affect this result [91, 121].

A brief aside: a P momentum state on an NS1 string with winding charge n_1 is split into "fractional units" of momentum: $\frac{P}{n_1}$, this will be an important feature which we will use below [121, 126].

What does our above 2-charge system look like in terms of D1-D5 branes, after using the above dualities? Just as P momentum states on n_1 NS1 strings are split into "fractional units" the D1 branes, after being bound to n_5 D5 branes, are split into fractional units of $\frac{1}{n_5}$ times their unbound values [121]. In Figure 5.7 we see that taking n_1 D1 branes (wrapped on S^1) and binding this to n_5 D5 branes (wrapped on $T^4 \times S^1$) produces an "effective" D1 brane with winding number $n_1 n_5$; this is due to the fact that the D1 brane wrapped on D5 has its tension (originating from its Planckian dimensions) reduced by a factor n_5 [121]. This multiplicative winding number $n_1 n_5$ is redolent of the one we found for the 2-charge entropy in (5.27), this factor reappearing is unsurprising as each case is equivalent after taking into account dualities; to be clear, P wrapped on NS1 is dual to D1 wrapped on D5, therefore the same physical results should emerge from studying either case.

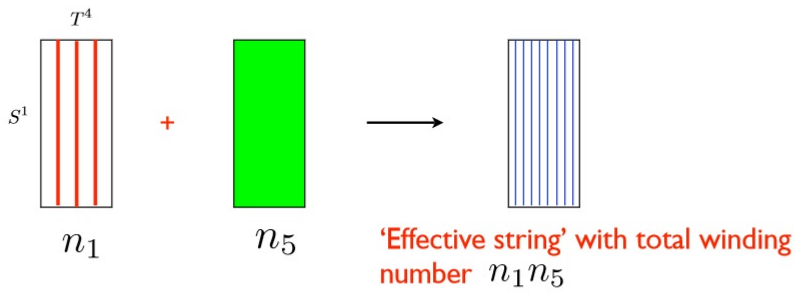

FIGURE 5.7: Pictorial view of the effect of binding n_1 D1 branes to n_5 D5 branes. Illustration from [121].

Adding momentum P to a D1-D5 bound state gives us a 3-charge system – see Figure 5.8 – with microscopic entropy

$$S_{micro} = 2\pi\sqrt{n_1 n_5 n_p} \tag{5.29}$$

where n_1 and n_5 are the number of D1 and D5 branes respectively, and n_p is the momentum charge of P; this result is derived by using supersymmetric arguments to account for the bosonic and fermionic modes, and by taking into account that the D1 branes are constrained by the D5 brane thus limiting their degrees of freedom [84, 121]. An RN black hole made in type IIB supergravity with the mass (and therefore also charge) of the BPS 3-charge system we just looked at gives a Bekenstein entropy equal to the value given in (5.29) [26], therefore the fuzzball microstate count for a 3-charge system produces the expected entropy of the associated black hole – a very suggestive result that fuzzballs may provide the hair required to resolve the long-standing entropy puzzle.

It would be nice to be able to determine the entropy of the more general non-BPS states. This can be done by allowing momentum modes P to run in opposite directions around the NS1 string, in this case the momentum charge cancels but the energy adds; therefore we get states in which the mass is greater than the charge, this has reproduced the correct entropy in the non-BPS 2-charge and 3-charge cases [121]. These P modes can also

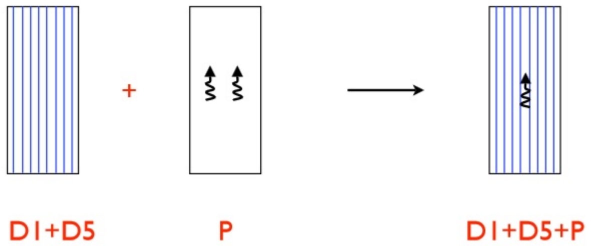

FIGURE 5.8: Pictorial view of D1-D5-P 3-charge system. Illustration from [121].

collide and "come off", exiting the system at a certain rate, the associated supergravity black holes emit Hawking radiation at the same rate, suggesting that the collision of momentum modes (or any phenomena related by dualities) present an alternative viewpoint of the Hawking emission mechanism [26, 121, 127, 128].

Do the brane-systems we looked at have a size of order the black hole horizon? This would explain the classical size of the black hole. The key to explaining the size of brane-systems is – as already mentioned – that branes bound to other branes have their tension reduced, thus giving a larger size than if they were unbound and had higher tension [121]. It turns out that the size of the brane-system grows with the number of charges (momentum, winding etc.) and in certain cases reproduces the Schwarzschild radius of the black hole [129].

The P modes on an NS1 string that we looked at above can only vibrate in transverse directions, not longitudinal, because the NS1 string is fundamental, it isn't made of smaller elements – these transverse vibrations create a non-zero spread in the non-compact directions [26, 121]. This spread turns out to be equal to the horizon size of the associated black hole [91]. If all the available energy is put in a few harmonics of the NS1 string then we get coherent states, if it's shared over many different harmonics then we get a generic quantum state (fuzzball) which has area of order microscopic entropy, suggesting that black holes may indeed be described by fuzzballs [91].

A collapsing shell of matter has a very small tunnelling amplitude to a fuzzball state however this is compensated by the very large amount of available fuzzball states – these factors cancel to give an order 1 probability for the matter to tunnel to a fuzzball state for times less than the Hawking evaporation time [121].

The fuzzball proposal is very persuasive as it removes the vacuum state at the black hole horizon, replacing it by a bound state of strings and branes. The horizon vacuum state – when stretched during hole evaporation – directly led to a mixed, information-less state whereas fuzzball radiation does not emerge from a vacuum region. It emerges from regions (microstate constructions) containing information about the system, therefore radiating in a way analogous to a macroscopic object [91, 103, 106].

In the future further research needs to be done on the dynamics of fuzzballs and a study into more general nonextremal microstate constructions, as well as a more quantitative look at the tunnelling of normal matter into fuzzball states [121]. This will illuminate the relation between fuzzballs and black holes and put this solution to the information paradox on firmer ground.

5.2.11 Supertranslations

Recently an idea was put forward that the information of a particle falling into a black hole is encoded on its horizon by "supertranslations" [130].

Supertranslations refers to the symmetries of future null infinity under the BMS group which describes asymptotic isometries of spacetime [131, 132]. The vacuum in general relativity is considered to be highly degenerate with the vacua related by supertranslations associated with a spontaneously broken BMS symmetry – the presence of radiation can induce transitions between these vacua [133]. An effect called "quantum memory" (shifts in detector positions and times due to nearby energy-momentum) can in theory be measured and has been shown to be equivalent to BMS transformations [133, 134]. A supertranslation can therefore shift spacetime coordinates at future null infinity and so if matter falling into a black hole radiates this can leave an imprint – a form of sought after hair cf. section 5.2.3 – at infinity as the vacuum there is shifted. This opens up the possibility that detectors positioned near future null infinity could measurably shift, recording the information about how the black hole formed [133].

Professor Stephen Hawking [130] has suggested that the horizons of stationary black holes also experience supertranslations, as the generators of the horizon are shifted by infalling particles. Further evidence for this can be found in *Polchinski* [135].

These supertranslations would manifest as delays in the emission of Hawking radiation by the hole – encoding the information about what fell in – therefore information is indeed conserved; the S-matrix is also invariant under these supertranslations and so there is no pure-to-mixed state evolution [130]. Further, this suggestion would be satisfying if shown to be correct as it would presumably remove the need for certain unfavoured concepts such as interiors of black holes leading to other universes and remnants.

The idea seems to avoid the problem of steadily increasing entanglement entropy between the interior and exterior of black holes during evaporation that we looked at earlier in section 5.1. The increasing entanglement result was predicated on the horizon being a vacuum – information-free – creating Hawking pairs with no relation to whatever formed the hole. Supertranslations may undermine this assumption by providing a horizon rich with information, which can then be carried out by the Hawking radiation thus solving the information paradox.

Applications of supertranslations to black hole horizons is still in its infancy and further work is due to be published on the subject in the coming months [130].

Chapter 6

Conclusion

In this dissertation we have set up the black hole information paradox and reviewed the most promising attempts so far at toppling it.

We began with a short look at the history of black holes as an idea and some of the most powerful mathematical tools that are commonly used to study them. A comparison of the rules that have been developed for describing black holes and the laws of classical thermodynamics was also included to show the striking similarities. We chose to derive Hawking radiation from black holes by way of the Unruh effect as we found this to be illuminating in showing how the concepts of particles and the vacuum are contingent on the state of observers in quantum field theory. Information and its relationship with entropy – specifically entanglement entropy – was introduced in order to clarify the main conceptual points involved in the paradox, before moving on to an appraisal of possible solutions.

The first potential solution we looked at involved remnants, a relatively old idea which suggested that holes stop evaporating before becoming Planck-sized where it was assumed quantum gravity effects would become important. This suffered from numerous problems as discussed, information emerging from black holes right at the end of evaporation led to some of the same issues. We also looked at a recent interesting paper treating black holes as a collection of Rubik's cubes, although this turned out only to be able to model the burning (and emission of radiation) of macroscopic objects, not including black holes. Another intriguing paper concerned the conjectured equivalence of entanglement and wormholes and its applicability to the information paradox, it is not clear at this time how applicable this equivalence will prove to be.

A very important event in the history of the paradox involves the AdS/CFT duality. This relationship convinced many physicists that information had to be conserved in all physical processes. Of course this didn't solve the problem but it did frame the paradox in a new way, most researchers now are convinced black holes evolve unitarily.

Several of the proposed solutions had unappealing elements which we discussed including: the infinite degeneracy of suggested Planck-sized remnants, the contradictions inherent in complementarity as shown by the AMPS paper authors, as well as the conclusion of the AMPS authors that the horizon of a black hole must be a highly energetic and lethal place for infalling observers – contradicting the well-established equivalence principle of general relativity.

In the author's opinion the most promising attempt at a solution to the paradox so far is given by the fuzzball, emerging from superstring theory. This has allowed the microscopic entropy of several types of holes to be accurately rederived by a count of bound states of strings and branes. Fuzzballs also provide a mechanism for unitary emission of radiation, and again the rate of emission agrees with that predicted by semiclassical theory. Further investigations into the tunnelling of collapsing matter into fuzzball states and their use in describing more general black holes will shed more light on these exotic objects and their link to semiclassical black holes.

The root cause of the paradox is the information-free vacuum state at the event horizon as predicted by general relativity. As black holes evaporate this vacuum "stretches" producing Hawking pairs independent of the matter that originally formed the hole. The recent proposal by Professor Stephen Hawking that supertranslations encode information at the horizon may remove this information-free property providing a mechanism for information emission. This idea is encouraging as it is not reliant on small corrections to the semiclassical derivation of Hawking radiation – which as we saw would be insufficient.

From our consideration of the current state of the paradox it seems as if research at this time is as active – if not more so – than at any time in the past, stimulated in large part by the discovery of the AMPS firewall. A resolution of the paradox, irrespective of whatever particular idea or framework provides the answer, would be a major step forward in our understanding of the universe and the next few years ought to be a very exciting time in this area of physics.

Bibliography

[1] J. Al-Khalili, *Paradox: The Nine Greatest Enigmas in Science*, Bantam Press [2012].

[2] R. d'Inverno, *Introducing Einstein's Relativity*, Clarendon Press [1992], pg 38.

[3] L. Susskind, J. Lindesay, *An Introduction To Black Holes, Information And The String Theory Revolution: The Holographic Universe*, World Scientific Publishing [2005].

[4] D. Topper, *How Einstein Created Relativity out of Physics and Astronomy* , Springer [2013], pgs 9-14.

[5] S.D. Mathur, *What Exactly is the Information Paradox?*, Lect. Notes Phys. **769** [2009], pgs 3-48, arXiv:0803.2030v1

[6] J. Michell, *On the means of discovering the distance, magnitude etc. of the fixed stars*, Philos. Trans. R. Soc. Lond. **74** [1783], pgs 35-57.

[7] J. Bradley, *Account of a new discovered motion of the fix'd stars*, Philos. Trans. R. Soc. Lond. **35** [1727], pgs 637-661.

[8] Colin Montgomery, Wayne Orchiston and Ian Whittingham, *Michell, Laplace and the origin of the black hole concept*, Journal of Astronomical History and Heritage, 12(2) [2009], pgs 90-96.

[9] S. W. Hawking (editor), *A Stubbornly Persistent Illusion: The Essential Scientific Works of Albert Einstein*, Perseus Books [2009], pg 308.

[10] P. Fré, V. Gorini, G. Magli, U. Moschella (editors), *Classical and Quantum Black Holes*, IOP Publishing [1999], Chapter 1.

[11] A. Fresnel, *The Wave Theory of Light - Memoirs by Huygens, Young and Fresnel*, American Book Company [1819].

[12] F. Dowker, *Black Holes*, Imperial College London, MSc Quantum Fields and Fundamental Forces, Lecture Notes [2013].

[13] Kip Thorne, *Black Holes and Time Warps: Einstein's Outrageous Legacy*, W. W. Norton and Company [1995], Chapter 3.

[14] S. Chandrasekhar, *The Maximum Mass of Ideal White Dwarfs*, Astrophys. J. **74** [1931], pgs 81-82.

[15] S. Carroll, *Spacetime and Geometry: An Introduction to General Relativity*, Pearson [2013].

[16] J. R. Oppenheimer, G. M. Volkoff, *On Massive Neutron Cores*, Physical Review **55** [1939], pgs 374-381.

[17] J.R. Oppenheimer, H. Snyder, *On Continued Gravitational Contraction*, Physical Review **56** [1939], pgs 455-459.

[18] R. Penrose, R. M. Floyd, *Extraction of Rotational Energy from a Black Hole*, Nature Physical Science **229** [1971], pgs 177-179

[19] R. Penrose, *The question of cosmic censorship*, Journal of Astrophysics and Astronomy **20** (3) [1999], pgs 233-248.

[20] J. M. Bardeen, B. Carter, S. W. Hawking, *The four laws of black hole mechanics*, Commun. Math. Phys. **31** [1973], pgs 161-170.

[21] J. D. Bekenstein, *Black holes and entropy*, Phys. Rev. D **7** [1973], pgs 2333-2346.

[22] S. W. Hawking, *Black hole explosions?*, Nature **248** [1974], pgs 30-31.

[23] T. Jacobson, *Introductory Lectures on Black Hole Thermodynamics*, Institute for Theoretical Physics, University of Utrecht [1996], available at physics.umd.edu/grt/taj/776b/lectures.pdf

[24] S. W. Hawking, *Particle creation by black holes*, Commun. Math. Phys. **43** [1975], pgs 199-220.

[25] S. W. Hawking, *Breakdown of predictability in gravitational collapse*, Phys. Rev. D **14** [1976], pg 2460.

[26] S. D. Mathur, *The fuzzball proposal for black holes: an elementary review*, Fortschritte der Physik **53** [2005], pgs 793-827, arXiv:hep-th/0502050v1

[27] R. Penrose, *Cycles of Time: An Extraordinary New View of the Universe*, Vintage [2011], pgs 184-192.

[28] L. Susskind, *The Black Hole War: My Battle with Stephen Hawking to Make the World Safe for Quantum Mechanics*; Little, Brown and Company [2008].

[29] J. Binney, D. Skinner, *The Physics of Quantum Mechanics*, OUP Oxford [2013], pgs 66-87.

[30] J. Preskill, *Do Black Holes Destroy Information?*, International Symposium on Black Holes, Membranes, Wormholes, and Superstrings; [1992], arXiv:hep-th/9209058

[31] J. Maldacena, *The Large N Limit of Superconformal Field Theories and Supergravity*, Adv. Theor. Math. Phys., **2** [1998], pgs 231-252.

[32] J. Maldacena, *The Illusion of Gravity*, Scientific American, **293** [2005], pgs 56-63.

[33] D. A. Lowe, L. Thorlacius, *AdS/CFT and the Information Paradox*, Phys. Rev. D. **60** [1999], arXiv:hep-th/9903237v2

[34] J. Maldacena, *Eternal black holes in Anti-de-Sitter*, JHEP **0304** [2003], arXiv:hep-th/0106112

[35] S. W. Hawking, *Information Loss in Black Holes*, Phys. Rev. D **72** [2005], arXiv:hep-th/0507171v2

[36] L. Susskind, L. Thorlacius, J. Uglum, *The Stretched Horizon and Black Hole Complementarity*, Phys. Rev. D. **48** [1993], pgs 3743-3761.

[37] A. Almheiri, D. Marolf, J. Polchinski, J. Sully, *Black Holes: Complementarity or Firewalls?*, JHEP **1302** [2013], arXiv:1207.3123.

[38] R. M. Wald, *General Relativity*, University of Chicago Press [1984].

[39] P.K. Townsend, *Black Holes*, Cambridge University Mathematical Tripos Part III, Black Holes Lecture Notes [1997], arXiv:gr-qc/9707012v1

[40] T. E. Regge, *General relativity without coordinates*, Nuovo Cim. **19** [1961], pgs 558–571.

[41] R. C. Henry, *Kretschmann Scalar for a Kerr-Newman Black Hole* [2009], arXiv:astro-ph/9912320v1

[42] G. Cruciani, *Deriving the Regge-Wheeler and Zerilli equations in the general static spherically symmetric case with Mathematica*^{TM}*and MathTensor*^{TM}, Nuovo Cim. B **115** [2000], pgs 693-698.

[43] M. Blau, *Lecture Notes on General Relativity*, University of Bern: Albert Einstein Center for Fundamental Physics [2014], available at www.blau.itp.unibe.ch/Lecturenotes.html

[44] E. W. Weisstein, *Hyperbolic Functions*, from MathWorld – A Wolfram Web Resource, available at mathworld.wolfram.com/HyperbolicFunctions.html

[45] E. Poisson, *A Relativist's Toolkit: The Mathematics of Black-Hole Mechanics*, Cambridge University Press [2007].

[46] T. Rowland, *Manifold*, from MathWorld – A Wolfram Web Resource, available at mathworld.wolfram.com/Manifold.html

[47] C. M. Hull, *Differential geometry*, Imperial College London, MSc Quantum Fields and Fundamental Forces, Lecture Notes [2014].

[48] E. W. Weisstein, *Killing Vectors*, from MathWorld – A Wolfram Web Resource, available at mathworld.wolfram.com/KillingVectors.html

[49] M. P. Hobson, G. P. Efstathiou, A. N. Lasenby, *General Relativity: An Introduction for Physicists*, Cambridge University Press [2006].

[50] S. Carroll, *General Relativity: Chapter 5 – More Geometry*, Caltech Lecture Notes [2014], available at ned.ipac.caltech.edu/level5/March01/Carroll3/Carroll5.html

[51] T. Lancaster, S.J. Blundell, *Quantum Field Theory for the Gifted Amateur*, OUP Oxford [2014].

[52] J. Polchinski, *String Theory: Volume 1, An Introduction to the Bosonic String*, Cambridge University Press [2005].

[53] J. Polchinski, *String Theory: Volume 2, Superstring Theory and Beyond*, Cambridge University Press [2005].

[54] T. Thiemann, *Lectures on Loop Quantum Gravity*, 271st WE Heraeus Seminar, arXiv:gr-qc/0210094v1

[55] R. M. Wald, *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*, University of Chicago Press [1994].

[56] T. Evans, *Quantum Field Theory*, Imperial College London, MSc Quantum Fields and Fundamental Forces, Lecture Notes [2015].

[57] M. H. S. Amin, *Consistency of the Adiabatic Theorem*, Phys. Rev. Lett. **102** [2009], arXiv:0810.4335

[58] D. S. Lemons, *A Student's Guide to Entropy*, Cambridge University Press [2013].

[59] C. Pickover, *From Archimedes to Hawking: Laws of Science and the Great Minds Behind Them*, OUP USA [2008], pg 328.

[60] S. J. Blundell, K. M. Blundell, *Concepts in Thermal Physics*, OUP Oxford [2009].

[61] R. V. L. Hartley, *Transmission of Information*, Bell System Technical Journal **7** (3) [1928], pgs 535-563.

[62] C. E. Shannon, *A Mathematical Theory of Communication*, Bell System Technical Journal **27** (3) [1948], pgs 379-423.

[63] M. M. Wilde, *Quantum Information Theory*, Cambridge University Press [2013].

[64] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, translation by R. T. Beyer, Princeton University Press [1955].

[65] S. J. van Enk, *Mixed states and pure states*, University of Oregon, Lecture Notes [2009].

[66] F. A. Bais, J. D. Farmer, *The Physics of Information*, Santa Fe Institute [2007], arXiv:0708.2837v2

[67] F. Rioux, *Elements of Dirac Notation*, St. John's University and College of St. Benedict, Lecture Notes, available at www.users.csbsju.edu/∼frioux/dirac/dirac.pdf

[68] T. Nishioka, S. Ryu, T. Takayanagi, *Holographic Entanglement Entropy: An Overview*, Journal of Physics A: Mathematical and Theoretical **42** (50) [2009], arxiv.org/abs/0905.0932v2

[69] M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press [2000].

[70] J. Martikainen, *Quantum Information and Computing – Lecture 13: Quantum Information*, University of Helsinki, Lecture Notes [2007], available at theory.physics.helsinki.fi/∼kvanttilaskenta/Lecture13.pdf

[71] M. S. Longair, *Theoretical Concepts in Physics: An Alternative View of Theoretical Reasoning in Physics*, Cambridge University Press [2003], pg 337.

[72] J. D. Bekenstein, *Black holes and information theory*, Contemporary Physics **45** (1) [2004], pgs 31-43, arXiv:quant-ph/0311049v1

[73] G. 't Hooft, *Dimensional Reduction in Quantum Gravity* [2009: newest version], arXiv:gr-qc/9310026v2

[74] L. Susskind, *The World as a Hologram*, Journal of Mathematical Physics **36** [1995], pgs 6377-6396.

[75] R. Bousso, *The holographic principle*, Reviews of Modern Physics **74** [2002], pgs 825-874.

[76] D. C. Robinson, *Four decades of black hole uniqueness theorems*, in *The Kerr Spacetime: Rotating Black Holes in General Relativity*, Cambridge University Press [2009], pgs 115-143.

[77] R. B. Mann. *Black Holes: Thermodynamics, Information, and Firewalls*, Springer [2015].

[78] V. P. Frolov, I. Novikov, *Dynamical origin of the entropy of a black hole*, Phys. Rev. D **48** [1993], pgs 4545-4551.

[79] W. Zurek, K. S. Thorne, *Statistical mechanical origin of the entropy of a rotating, charged black hole*, Phys. Rev. Lett. **54** [1985], pgs 2171-2175.

[80] G. 't Hooft, *On the quantum structure of a black hole*, Nucl. Phys. B **256** [1985], pgs 727-745.

[81] L. Bombelli, R. K. Koul, J. Lee, R. D. Sorkin, *A quantum source of entropy for black holes*, Phys. Rev. D **34** [1986], pgs 373-383.

[82] M. Srednicki, *Entropy and area*, Phys. Rev. Lett. **71** [1993], pgs 666-669.

[83] R. M. Wald, *Black hole entropy is the Noether charge*, Phys. Rev. D **48** [1993], pgs 3427-3431.

[84] A. Strominger, C. Vafa, *Microscopic origin of the Bekenstein-Hawking entropy*, Phys. Lett. B **379** [1996], pgs 99-104.

[85] C. Rovelli, L. Smolin, *Loop space representation of quantum general relativity*, Nucl. Phys. B. **331** [1990], pgs 80-152.

[86] A. Ashtekar, J. Baez, A. Corichi, K. Krasnov, *Quantum geometry and black hole entropy*, Phys. Rev. Lett. **80** [1998], pgs 904-907.

[87] S. D. Mathur, *The information paradox: a pedagogical introduction*, Class. Quantum Grav. **26** [2009], available at stacks.iop.org/CQG/26/224001

[88] J. Preskill, *Black holes and information: A crisis in quantum physics*, Caltech Theory Seminar [1994], available at theory.caltech.edu/∼preskill/talks/blackholes.ps

[89] S. D. Mathur, *What the information paradox is not* [2011], arXiv:1108.0302v2

[90] P. Chen, Y. C. Ong, D. Yeom, *Black Hole Remnants and the Information Loss Paradox* [2015], arXiv:1412.8366v2

[91] S. D. Mathur, *The fuzzball paradigm for black holes: FAQ* [2009], available at physics.ohio-state.edu/∼mathur/faq2.pdf

[92] S. Coleman, J. Preskill, F. Wilczek, *Quantum hair on black holes*, Nucl. Phys. B **378** [1992], pgs 175-246.

[93] F. Larsen, F. Wilczek, *Classical Hair in String Theory I: General Formulation*, Nucl. Phys. B **475** [1996], pgs 627-644.

[94] S. Coleman, *Why there is nothing rather than something: a theory of the cosmological constant*, Nucl. Phys. B **310** [1988], pgs 643-668.

[95] S. W. Hawking, *Baby Universes II*, Mod. Phys. Lett. A **5** [1990], pgs 453-466.

[96] R. D. Carlitz, R. S. Willey, *Lifetime of a black hole*, Phys. Rev. D **36** [1987], pgs 2336-2341.

[97] J. A. Smolin, J. Oppenheim, *Locking information in black holes*, Phys. Rev. Lett. **96** [2006], arXiv:hep-th/0507287v2

[98] B. Czech, K. Larjo, M. Rozali, *Black Holes as Rubik's Cubes* [2011], arXiv:1106.5229v1

[99] M. K. Parikh, F. Wilczek, *Hawking Radiation As Tunneling*, Phys. Rev. Lett. **85** [2000], pgs 5042-5045.

[100] L. Vanzo, G. Acquavivam, R. D. Criscienzo, *Tunnelling methods and Hawking's radiation: achievements and prospects*, Class. Quantum Grav. **28** [2011], available at stacks.iop.org/CQG/28/183001

[101] R. Kerner, R. B. Mann, *Fermions tunnelling from black holes*, Class. Quantum Grav. **25** [2008], 17 pgs.

[102] C. R. Stephens, G. 't Hooft, B. F. Whiting, *Black hole evaporation without information loss*, Class. Quantum Grav. **11** [1994], pgs 621-648.

[103] S. D. Mathur, *Black holes and beyond*, Annals of Physics **327** [2012], pgs 2760-2793, arXiv:1205.0776v2

[104] J. Polchinski, *Review: Black Holes Information, Entanglement, Complementarity, and Firewalls*, Simons Symposium [2013], available at qpt.physics.harvard.edu/simons

[105] D. N. Page, *Information in black hole radiation*, Phys. Rev. Lett. **23** [1993], pgs 3743-3746.

[106] S. D. Mathur, D. Turton, *The flaw in the firewall argument*, Nucl. Phys. B **884** [2014] pgs 566-611, arXiv:1306.5488v3

[107] Z. Merali, *Astrophysics: Fire in the hole!*, Nature **496** [2013], pgs 20-23.

[108] D. N. Page, *Time dependence of Hawking radiation entropy*, JCAP **1309** [2013], arXiv:1301.4995v3

[109] J. Maldacena, L. Susskind, *Cool horizons for entangled black holes*, Fortschritte der Physik **61** [2013], pgs 781-811, arXiv:1306.0533v2

[110] K. C. Cole, *Wormholes Untangle a Black Hole Paradox*, Quanta Magazine [2015], available at quantamagazine.org/20150424-wormholes-entanglement-firewalls-er-epr/

[111] J. Preskill, *Quantum Entanglement*, Caltech Lecture Notes [2001], available at theory.caltech.edu/∼preskill/ph229/notes/chap4_01.pdf

[112] D. Overbye, *A Black Hole Mystery Wrapped in a Firewall Paradox*, New York Times [2013], available at nytimes.com/2013/08/13/science/space/a-black-hole-mystery-wrapped-in-a-firewall-paradox.html

[113] J. de Boer, *Introduction to the AdS/CFT Correspondence*, University of Amsterdam, Institute for Theoretical Physics, Technical Report [2003].

[114] S. D. Mathur, *Confusions and questions about the information paradox*, available at www.physics.ohio-state.edu/∼mathur/confusions2.pdf

[115] I. Bengtsson, *Anti-de Sitter space*, Lecture Notes [1998], available at www.fysik.su.se/∼ingemar/Kurs.pdf

[116] A. V. Ramallo, *Introduction to the AdS/CFT correspondence* [2013], arXiv:1310.4319v3

[117] Y. Nakayama, *Scale invariance vs conformal invariance* [2014], arXiv:1302.0884v4

[118] V. Riva, J. Cardy, *Scale and conformal invariance in field theory: a physical counterexample* [2005], arXiv:hep-th/0504197v2

[119] L. Wassink, *N=4 Super Yang-Mills Theory*, University of Iowa, Presentation [2009], available at www-hep.physics.uiowa.edu/∼vincent/courses/29276/Presentations/Luke

[120] A. Rajantie, *Advanced Quantum Field Theory*, Imperial College London, MSc Quantum Fields and Fundamental Forces, Lecture Notes [2015].

[121] S. D. Mathur, *The black hole information problem and the fuzzball proposal*, CERN Winter School on Supergravity, Strings and Gauge Theories, Lecture Series (four in total) [2009], slides and videos available at indico.cern.ch/event/37753/overview

[122] O. Lunin, S. D. Mathur, *AdS/CFT duality and the black hole information paradox*, Nucl. Phys. B **623** [2002], pgs 342-394.

[123] O. Lunin, S.D. Mathur, *Statistical interpretation of Bekenstein entropy for systems with a stretched horizon*, Phys. Rev. Lett., **88** [2002], pgs 211303-1 to 211303-4.

[124] L. Susskind, *Some speculations about black hole entropy in string theory* [1993], arXiv:hep-th/9309145v2

[125] C. Pope, *Lectures on Kaluza-Klein Theory*, available at people.physics.tamu.edu/pope/ihplec.pdf

[126] S. D. Mathur, *The information paradox: conflicts and resolutions* [2012], arXiv:1201.2079v1

[127] S. R. Das, S. D. Mathur, *Comparing decay rates for black holes and D-branes*, Nucl. Phys. B **478** [1996], pgs 561-576.

[128] J. Maldacena, A. Strominger, *Black Hole Greybody Factors and D-Brane Spectroscopy*, Phys. Rev. D **55** [1997], pgs 861-870.

[129] S. D. Mathur, *Emission rates, the correspondence principle and the information paradox*, Nucl. Phys. B **529** [1998], pgs 295-320.

[130] S. W. Hawking, *The Information Paradox for Black Holes.* [2015], arXiv:1509.01147v1

[131] H. Bondi, M. G. van der Burg, A. W. Metzner, *Gravitational Waves in General Relativity. 7. Waves from Axi-symmetric Isolated Systems*, Proc. Roy. Soc. Lond. A **269** [1962], pg 21.

[132] R. K. Sachs, *Gravitational Waves in General Relativity. 8. Waves in Asymptotically Flat Space-time*, Proc. Roy. Soc. Lond. A **270** [1962], pg 103.

[133] A. Strominger, A. Zhiboedov, *Gravitational Memory, BMS Supertranslations and Soft Theorems* [2014], arXiv:1411.5745v1

[134] A. Tolish, L. Bieri, D. Garfinkle, R. M. Wald, *Examination of a simple example of gravitational wave memory*, Phys. Rev. D **90** [2014], arXiv:1405.6396v1

[135] J. Polchinski, *Chaos in the black hole S-matrix*, arXiv:1505.08108v2